

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 20:06:06 : Search time 4552.9 Seconds

(without alignments)
1021.811 Million cell updates/sec

Title: US-09-516-493-11

Perfect score: 282

Sequence: 1 gagcctgctgatttcacct.....tcttcagcttcctttcacg 282

Scoring table: OLIGO_NDC

Gapop 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database: GenEmbl:

1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_cm:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_rod:*
33: em_hlg_hum:*
34: em_hlg_inv:*
35: em_hlg_rod:*
36: em_hlg_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	282	100.0	2087	6 AX076669	AX076669 Sequence
2	282	100.0	2087	10 RNO245935	AJ245935 Rattus no
3	282	100.0	2189	10 AB033418	AB033418 Rattus no
4	50	17.7	1490	10 MM017802	Y17802 Mus musculu
5	50	17.7	1843	10 AF232061	AF232061 Mus muscu
6	50	17.7	2072	6 AX076671	AX076671 Sequence
7	50	17.7	2072	10 MM0245936	AJ245936 Mus muscu
8	44	15.6	1012	4 AF321324	AF321324 Bos tauru
9	41	14.5	1508	9 HSA17801	Y17801 Homo sapien
10	41	14.5	1873	9 HSA245937	AJ245937 Homo sapi
11	41	14.5	2080	6 AX179740	AX179740 Sequence
12	41	14.5	2217	6 AX076667	AX076667 Sequence
13	41	14.5	225370	9 AL445222	AL445222 Human DNA
14	23	8.2	1461	6 AX191507	AX191507 Sequence
15	21	7.4	11931	10 CGR00D	X61084 C.griusens r
16	21	7.4	12544	10 M05SLPSEXA	M73820 Mouse sex -1
17	21	7.4	17251	10 AB015623	AB015623 Mus muscu
18	21	7.4	26307	10 M05SLPSEXB	M64933 Mouse nonfu
19	21	7.4	149886	21 MMHC438N12	AF049850 Mus muscu
20	21	7.4	201964	10 MMHC29N7	AF030001 Mus muscu
21	21	7.4	201986	10 AC006289	AC006289 Mus muscu
22	21	7.4	239339	2 AC091703	AC091703 Mus muscu
23	20	7.1	7954	10 AF049340	AF049340 Mus muscu
24	20	7.1	25185	2 AL139799	AL139799 Homo sapi
25	20	7.1	35891	2 AC087711	AC087711 Homo sapi
26	20	7.1	37476	9 HS316G12	AL031709 Human DNA
27	20	7.1	41306	9 AP000523	AP000523 Homo sapi
28	20	7.1	98471	8 F14023	AC012654 Arabidops
29	20	7.1	110000	2 AL512625-2	Continuation (3 of
30	20	7.1	136630	2 AC023572	AC023572 Homo sapi
31	20	7.1	148228	2 AC051659	AC051659 Homo sapi
32	20	7.1	151418	2 AC025100	AC025100 Homo sapi
33	20	7.1	162208	9 AC073321	AC073321 Homo sapi
34	20	7.1	164485	2 AC026106	AC026106 Homo sapi
35	20	7.1	167320	2 AF298853	AF298853 Homo sapi
36	20	7.1	170340	2 AC092703	AC092703 Homo sapi
37	20	7.1	175271	9 AL354718	AL354718 Human DNA
38	20	7.1	176845	9 AL162252	AL162252 Human DNA
39	20	7.1	180669	2 CNS07EG4	AL588743 Homo sapi
40	20	7.1	182817	2 AC080181	AC080181 Homo sapi
41	20	7.1	209643	10 AC080018	AC080018 Mus muscu
42	20	7.1	240000	2 AC008130	AC008130 Homo sapi
43	20	7.1	279011	9 AE006467	AE006467 Homo sapi
44	19	6.7	800	4 AY033976	AY033976 Epesiscus
45	19	6.7	1722	10 MMRALDH2	X99273 M.musculus
46	19	6.7	2240	10 RND60063	U60063 Rattus norv
47	19	6.7	2707	9 BC001060	BC001060 Homo sapi
48	19	6.7	2711	9 HSPAX8A	X69699 H.sapiens P
49	19	6.7	4575	9 HSMYBPC	X84075 H.sapiens m
50	19	6.7	10697	1 AE007445	AE007445 Streptoco
51	19	6.7	20662	9 HSMYBPC3	Y10129 H.sapiens m
52	19	6.7	24066	9 HSU91629	U91629 Human card
53	19	6.7	98743	9 AL158172	AL158172 Human DNA
54	19	6.7	136555	2 AC068279	AC068279 Homo sapi
55	19	6.7	150052	2 AL353779	AL353779 Homo sapi
56	19	6.7	154630	2 AC018410	AC018410 Homo sapi
57	19	6.7	157750	2 AC006097	AC006097 Homo sapi
58	19	6.7	157868	2 AC080043	AC080043 Homo sapi
59	19	6.7	161630	2 AC024704	AC024704 Homo sapi
60	19	6.7	161706	2 AC027622	AC027622 Homo sapi
61	19	6.7	170749	9 AC025272	AC025272 Homo sapi
62	19	6.7	177427	9 CNS018P1	AL110292 Human chr
63	19	6.7	179937	9 AC016683	AC016683 Homo sapi
64	19	6.7	183536	2 AC019059	AC019059 Homo sapi
65	19	6.7	185231	2 AC023566	AC023566 Homo sapi
66	19	6.7	187109	2 AC074100	AC074100 Homo sapi
67	19	6.7	188349	2 AL390197	AL390197 Homo sapi
68	19	6.7	188459	2 AC090582	AC090582 Homo sapi
69	19	6.7	192202	2 AL133482	AL133482 Homo sapi
70	19	6.7	206784	9 HS9317	AL022401 Homo sapi

SUMMARIES

8

C 71	19	6.7	219574	9	AC016759	Homo sapi	144	18	6.4	174811	2	AC013259	Homo sapi
C 72	19	6.7	297172	2	SPNEU905	AL449927 Streptoco	C 145	18	6.4	174875	2	AC024682	Homo sapi
C 73	18	6.4	585	11	664711	6449711 B12B19 GSS	C 146	18	6.4	175707	2	AC022030	Homo sapi
C 74	18	6.4	2838	10	RNU87306	U87306 Rattus norv	C 147	18	6.4	177169	9	CNS01RGE	Human chr
C 75	18	6.4	5092	8	FSU61838	U61838 Fusarium so	C 148	18	6.4	178191	2	AP001120	Homo sapi
C 76	18	6.4	10142	1	AE004957	AE004957 Pseudomon	C 149	18	6.4	178733	2	AF238377	Homo sapi
C 77	18	6.4	10442	6	AX060713	AX060713 Sequence	C 150	18	6.4	181485	2	AC024036	Homo sapi
C 78	18	6.4	10442	6	AX060892	AX060892 Sequence	C 151	18	6.4	181499	2	AC074237	Homo sapi
C 79	18	6.4	10442	6	AF285167	AF285167 Homo sapi	C 152	18	6.4	181640	2	AC091020	Homo sapi
C 80	18	6.4	10474	6	AX060719	AX060719 Sequence	C 153	18	6.4	183412	2	AP001899	Homo sapi
C 81	18	6.4	10474	6	AX060721	AX060721 Sequence	C 154	18	6.4	184210	2	AC011582	Homo sapi
C 82	18	6.4	10474	6	AX060898	AX060898 Sequence	C 155	18	6.4	185967	9	AC012508	Homo sapi
C 83	18	6.4	10474	6	AX060900	AX060900 Sequence	C 156	18	6.4	186526	2	AC026495	Homo sapi
C 84	18	6.4	19089	2	AC026872	AC026872 Homo sapi	C 157	18	6.4	186673	2	AL590713	Homo sapi
C 85	18	6.4	23600	3	LMFL5856	AL537592 Leishmani	C 158	18	6.4	187404	2	AP003545	Oryza sat
C 86	18	6.4	41796	2	AC090489	AC090489 Mus muscu	C 159	18	6.4	187404	2	AP003545	Oryza sat
C 87	18	6.4	45985	2	AC020762	AC020762 Homo sapi	C 160	18	6.4	187984	2	AP001788	Homo sapi
C 88	18	6.4	52947	2	AC068354	AC068354 Homo sapi	C 161	18	6.4	188888	2	AC027142	Homo sapi
C 89	18	6.4	70311	9	AC002390	AC002390 Human DNA	C 162	18	6.4	188891	2	AC021156	Homo sapi
C 90	18	6.4	74155	9	AC004656	AC004656 Homo sapi	C 163	18	6.4	189604	2	AC079459	Homo sapi
C 91	18	6.4	86162	9	HSJ988G17	AL078643 Human DNA	C 164	18	6.4	190982	2	AC025106	Homo sapi
C 92	18	6.4	90113	8	AC005956	AC005956 Arabidops	C 165	18	6.4	191337	2	AC026711	Homo sapi
C 93	18	6.4	91303	8	ATFE2413	AL138655 Arabidops	C 166	18	6.4	193179	2	AC027084	Homo sapi
C 94	18	6.4	98572	2	AC024573	AC024573 Homo sapi	C 167	18	6.4	193731	9	HS80N2	Human DNA
C 95	18	6.4	99558	2	AF186596	AF186596 Homo sapi	C 168	18	6.4	194995	2	AC018696	Homo sapi
C 96	18	6.4	108893	9	HSBDJ130E4	AL078582 Human DNA	C 169	18	6.4	195383	2	AC011804	Homo sapi
C 97	18	6.4	115045	9	AC004414	AC004414 Homo sapi	C 170	18	6.4	195635	2	AC017101	Homo sapi
C 98	18	6.4	123288	9	AP000547	AP000547 Homo sapi	C 171	18	6.4	199853	2	AC073785	Homo sapi
C 99	18	6.4	125666	2	AC092080	AC092080 Homo sapi	C 172	18	6.4	204231	2	CNS00M8S	Mus muscu
C 100	18	6.4	129290	2	AC020761	AC020761 Homo sapi	C 173	18	6.4	207392	2	AC006558	Homo sapi
C 101	18	6.4	132700	2	AC040939	AC040939 Homo sapi	C 174	18	6.4	214125	2	AC020907	Homo sapi
C 102	18	6.4	137103	2	AL353593	AL353593 Homo sapi	C 175	18	6.4	214292	2	AC026894	Homo sapi
C 103	18	6.4	138858	8	AP002968	AP002968 Oryza sat	C 176	18	6.4	215734	2	AC073710	Homo sapi
C 104	18	6.4	139376	9	HSU95742	U95742 Human: chrom	C 177	18	6.4	217000	1	AP000004	Mus muscu
C 105	18	6.4	141208	2	AP001492	AP001492 Homo sapi	C 178	18	6.4	223235	2	AC020839	Mus muscu
C 106	18	6.4	143609	2	AC068816	AC068816 Homo sapi	C 179	18	6.4	233752	2	AC073683	Mus muscu
C 107	18	6.4	146391	8	AC074354	AC074354 Genomic S	C 180	18	6.4	243853	2	AC007601	Mus muscu
C 108	18	6.4	146867	2	AP003520	AP003520 Oryza sat	C 181	18	6.4	243853	2	AC007601	Mus muscu
C 109	18	6.4	147021	2	AC091321	AC091321 Homo sapi	C 182	17	6.0	275434	9	HS4A00879	Human chr
C 110	18	6.4	147029	9	AL137068	AL137068 Human DNA	C 183	17	6.0	360	9	AF103063	Homo sapi
C 111	18	6.4	147812	2	AC062025	AC062025 Homo sapi	C 184	17	6.0	365	11	HSPE28D01	Homo sapi
C 112	18	6.4	150159	2	AC018791	AC018791 Homo sapi	C 185	17	6.0	393	6	AX069719	Sequence
C 113	18	6.4	151041	8	AP0033258	AP0033258 Oryza sat	C 186	17	6.0	407	11	AF157573	Drosophila
C 114	18	6.4	156393	8	AP003204	AP003204 Oryza sat	C 187	17	6.0	421	6	AX069788	Sequence
C 115	18	6.4	157269	2	AP003096	AP003096 Homo sapi	C 188	17	6.0	616	11	G52782	SHGC-81669
C 116	18	6.4	157291	2	AC009091	AC009091 Homo sapi	C 189	17	6.0	801	10	AF346413	Mus muscu
C 117	18	6.4	158211	2	AC026853	AC026853 Homo sapi	C 190	17	6.0	879	6	AX144091	Sequence
C 118	18	6.4	158871	9	AL354833	AL354833 Human DNA	C 191	17	6.0	879	6	AX144311	Sequence
C 119	18	6.4	159015	2	AC091053	AC091053 Homo sapi	C 192	17	6.0	1215	5	AF327410	Danio rer
C 120	18	6.4	160004	2	AC013607	AC013607 Homo sapi	C 193	17	6.0	1497	6	AX135130	Sequence
C 121	18	6.4	160653	2	AC084192	AC084192 Homo sapi	C 194	17	6.0	1735	9	AF273270	Homo sapi
C 122	18	6.4	160737	2	AC024486	AC024486 Homo sapi	C 195	17	6.0	1767	9	AK001034	Homo sapi
C 123	18	6.4	160999	2	AP003369	AP003369 Oryza sat	C 196	17	6.0	1795	9	AF273271	Homo sapi
C 124	18	6.4	161866	2	AC010614	AC010614 Homo sapi	C 197	17	6.0	1801	8	AF139539	Raphanus
C 125	18	6.4	162930	2	AC016984	AC016984 Mus muscu	C 198	17	6.0	1815	9	AK000461	Homo sapi
C 126	18	6.4	163531	2	AC091483	AC091483 Trypanoso	C 199	17	6.0	1824	8	AF139538	Raphanus
C 127	18	6.4	163844	9	AC009048	AC009048 Homo sapi	C 200	17	6.0	2485	9	AF119045	Homo sapi
C 128	18	6.4	163939	2	AC079080	AC079080 Homo sapi	C 201	17	6.0	2687	5	AB035315	Paracalt
C 129	18	6.4	164485	2	AC026106	AC026106 Homo sapi	C 202	17	6.0	2848	9	AK023113	Homo sapi
C 130	18	6.4	164766	2	AC026657	AC026657 Homo sapi	C 203	17	6.0	2881	10	AY026945	Rattus no
C 131	18	6.4	164815	2	AL160252	AL160252 Homo sapi	C 204	17	6.0	2913	1	AF269866	Staphyloc
C 132	18	6.4	165202	2	AP003491	AP003491 Oryza sat	C 205	17	6.0	2913	6	AX145184	Sequence
C 133	18	6.4	165302	9	AC087237	AC087237 Homo sapi	C 206	17	6.0	3150	1	AF269656	Staphyloc
C 134	18	6.4	167288	8	AC084831	AC084831 Oryza sat	C 207	17	6.0	3150	6	AX144976	Sequence
C 135	18	6.4	167449	2	AC025073	AC025073 Homo sapi	C 208	17	6.0	3452	6	AX135128	Sequence
C 136	18	6.4	168837	9	AC090108	AC090108 Homo sapi	C 209	17	6.0	3514	9	AF084830	Homo sapi
C 137	18	6.4	168908	2	CNS07EE0	AL445883 Homo sapi	C 210	17	6.0	3589	1	AF269311	Staphyloc
C 138	18	6.4	170832	2	AC009454	AC009454 Homo sapi	C 211	17	6.0	3589	6	AX144631	Sequence
C 139	18	6.4	171390	2	AC015473	AC015473 Homo sapi	C 212	17	6.0	3938	9	HSM802861	Homo sapi
C 140	18	6.4	172538	2	AC041016	AC041016 Homo sapi	C 213	17	6.0	4198	1	AF269562	Staphyloc
C 141	18	6.4	172827	9	AC007216	AC007216 Homo sapi	C 214	17	6.0	4198	6	AX144882	Sequence
C 142	18	6.4	173996	2	AC025084	AC025084 Homo sapi	C 215	17	6.0	4360	6	AX180877	Sequence
C 143	18	6.4	174612	2	AC023889	AC023889 Homo sapi	C 216	17	6.0	7944	3	AF306784	Cherax qu

C 217	17	6.0	10261	1	AE001882	AE001882 Deinococc	290	17	6.0	137506	9	AC004824	AC004824 Homo sapi
C 218	17	6.0	10293	1	AE005012	AE005012 Halobacter	291	17	6.0	137947	9	HS4514	AL023581 Human DNA
C 219	17	6.0	17019	4	MTOAMTDNA	X83427 Ornithobact	292	17	6.0	137995	2	AC025009	AC025009 Homo sapi
C 220	17	6.0	18360	1	CHORF	X93204 C. burnetii	293	17	6.0	139006	9	AL162431	AL162431 Human DNA
C 221	17	6.0	22663	3	DROMHC	M61229 D. melanoagas	294	17	6.0	140137	2	AC027611	AC027611 Homo sapi
C 222	17	6.0	25076	9	AF039904	AF039904 Homo sapi	295	17	6.0	140168	2	AC021783	AC021783 Homo sapi
C 223	17	6.0	25676	9	AL451055	AL451055 Human DNA	296	17	6.0	141263	9	AC003687	AC003687 Homo sapi
C 224	17	6.0	29352	3	DMC86E4	AL021086 Drosophill	297	17	6.0	142008	9	AC025257	AC025257 Homo sapi
C 225	17	6.0	29839	9	HS133C6	2581082 Human DNA	298	17	6.0	142971	9	AC020629	AC020629 Homo sapi
C 226	17	6.0	32601	1	AF131076	AF131076 Coxiiella	299	17	6.0	143044	9	AC008750	AC008750 Homo sapi
C 227	17	6.0	33606	2	AC017861	AC017861 Drosophill	300	17	6.0	144474	9	AF051934	AF051934 Homo sapi
C 228	17	6.0	34419	2	AC023428	AC023428 Mus muscu	301	17	6.0	144707	2	AC084715	AC084715 Homo sapi
C 229	17	6.0	34725	1	SC1B2	AL356812 Streptolomy	302	17	6.0	145111	2	AL390245	AL390245 Homo sapi
C 230	17	6.0	34820	7	AF085222	AF085222 Streptococ	303	17	6.0	145347	2	AC073821	AC073821 Mus muscu
C 231	17	6.0	35437	1	SCF43A	AL096837 Streptolomy	304	17	6.0	145491	2	AC008193	AC008193 Drosophill
C 232	17	6.0	37329	1	CHOPH1	X75356 C. burnetii	305	17	6.0	145568	9	AC010000	AC010000 Homo sapi
C 233	17	6.0	38399	9	AP001477	AP001477 Homo sapi	306	17	6.0	145863	2	AC090255	AC090255 Homo sapi
C 234	17	6.0	38476	9	U73024	U73024 Homo sapien	307	17	6.0	146054	2	AC040963	AC040963 Homo sapi
C 235	17	6.0	39280	1	CHOPRS	Y15898 Coxiiella bu	308	17	6.0	146755	2	AC083982	AC083982 Homo sapi
C 236	17	6.0	40874	9	AF129108	AF129108 Homo sapi	309	17	6.0	148610	9	AC023426	AC023426 Homo sapi
C 237	17	6.0	41025	9	AC003975	AC003975 Human Cos	310	17	6.0	148849	9	AL158837	AL158837 Human DNA
C 238	17	6.0	42883	9	AC006167	AC006167 Homo sapi	311	17	6.0	149896	2	AC026467	AC026467 Homo sapi
C 239	17	6.0	47042	9	AC007031	AC007031 Homo sapi	312	17	6.0	150732	2	AC092413	AC092413 Felis cat
C 240	17	6.0	48287	2	AC017774	AC017774 Drosophill	313	17	6.0	151673	10	AC087102	AC087102 Rattus no
C 241	17	6.0	56131	9	AP000295	AP000295 Homo sapi	314	17	6.0	151704	2	AC026636	AC026636 Homo sapi
C 242	17	6.0	59315	2	AC084169	AC084169 Homo sapi	315	17	6.0	152116	2	AC048367	AC048367 Homo sapi
C 243	17	6.0	60057	9	AC023644	AC023644 Homo sapi	316	17	6.0	152318	2	AC093279	AC093279 Homo sapi
C 244	17	6.0	60144	2	AC019798	AC019798 Drosophill	317	17	6.0	153225	2	AC013581	AC013581 Homo sapi
C 245	17	6.0	61170	2	AC068163	AC068163 Homo sapi	318	17	6.0	153724	2	AC041034	AC041034 Homo sapi
C 246	17	6.0	63324	10	AC084043	AC084043 Mus muscu	319	17	6.0	153728	2	AP003355	AP003355 Homo sapi
C 247	17	6.0	63585	2	AC026983	AC026983 Homo sapi	320	17	6.0	153936	2	AC026636	AC026636 Homo sapi
C 248	17	6.0	65241	3	AC084458	AC084458 Caenorhab	321	17	6.0	154236	2	AL153529	AL153529 Homo sapi
C 249	17	6.0	68379	3	AC005119	AC005119 Drosophill	322	17	6.0	154272	2	AP001888	AP001888 Homo sapi
C 250	17	6.0	68478	8	NCB12E1	AL350091 Neurospor	323	17	6.0	154318	9	AC026791	AC026791 Homo sapi
C 251	17	6.0	69266	9	AL589882	AL589882 Human DNA	324	17	6.0	154685	9	AC007214	AC007214 Pan trogl
C 252	17	6.0	72060	9	AL592217	AL592217 Human DNA	325	17	6.0	155334	2	AC011156	AC011156 Homo sapi
C 253	17	6.0	72747	2	AC079582	AC079582 Mus muscu	326	17	6.0	155842	2	AL590085	AL590085 Homo sapi
C 254	17	6.0	72847	2	AC023563	AC023563 Homo sapi	327	17	6.0	155902	2	AP002852	AP002852 Homo sapi
C 255	17	6.0	77694	9	AC037461	AC037461 Homo sapi	328	17	6.0	156069	9	AC079256	AC079256 Homo sapi
C 256	17	6.0	83057	9	AL445703	AL445703 Human DNA	329	17	6.0	156264	2	AC0979329	AC0979329 Homo sapi
C 257	17	6.0	85089	9	AP001432	AP001432 Homo sapi	330	17	6.0	156457	2	AC093014	AC093014 Homo sapi
C 258	17	6.0	89994	10	AF332860S2	AF332861 Mus muscu	331	17	6.0	156492	2	AC020582	AC020582 Homo sapi
C 259	17	6.0	91228	8	ATF14M19	AL049480 Arabidops	332	17	6.0	157548	2	AC021972	AC021972 Homo sapi
C 260	17	6.0	91916	8	AC005964	AC005964 Arabidops	333	17	6.0	158034	2	AC019275	AC019275 Homo sapi
C 261	17	6.0	92884	2	AC092140	AC092140 Homo sapi	334	17	6.0	158267	9	AC053480	AC053480 Homo sapi
C 262	17	6.0	100000	9	AP000010	AP000010 Homo sapi	335	17	6.0	158357	9	AL365434	AL365434 Human DNA
C 263	17	6.0	100000	9	AP000044	AP000044 Homo sapi	336	17	6.0	158794	2	AC023961	AC023961 Homo sapi
C 264	17	6.0	100000	9	AP000112	AP000112 Homo sapi	337	17	6.0	158892	2	CNS01R14	AL162871 Homo sapi
C 265	17	6.0	100000	9	AP000151	AP000151 Homo sapi	338	17	6.0	159229	9	AL359180	AL359180 Human DNA
C 266	17	6.0	100000	2	AC090188	AC090188 Homo sapi	339	17	6.0	159251	2	AC021659	AC021659 Homo sapi
C 267	17	6.0	100141	2	AC091426	AC091426 Mus muscu	340	17	6.0	159299	2	AC091890	AC091890 Homo sapi
C 268	17	6.0	101269	9	AC005300	AC005300 Homo sapi	341	17	6.0	159766	9	AC010388	AC010388 Homo sapi
C 269	17	6.0	101841	9	AL451185	AL451185 Human DNA	342	17	6.0	160270	33	AC025087	AC025087 Homo sapi
C 270	17	6.0	103353	8	AC010704	AC010704 Arabidops	343	17	6.0	160350	2	AC074094	AC074094 Homo sapi
C 271	17	6.0	109310	9	AL133260	AL133260 Human DNA	344	17	6.0	161185	2	AP003571	AP003571 Oryza sat
C 272	17	6.0	110000	2	AL457133_2	Continuation (3 of	345	17	6.0	161590	2	AC092727	AC092727 Bos tauru
C 273	17	6.0	110000	2	AL442224_2	Continuation (3 of	346	17	6.0	162104	2	AC067976	AC067976 Homo sapi
C 274	17	6.0	110000	2	LMFLCHR36_26	Continuation (27 of	347	17	6.0	162507	9	AC090944	AC090944 Homo sapi
C 275	17	6.0	112018	9	AB023050	AB023050 Homo sapi	348	17	6.0	162560	2	AC069005	AC069005 Homo sapi
C 276	17	6.0	113196	9	HSJ697K14	AL121829 Human DNA	349	17	6.0	162601	9	AL603629	AL603629 Homo sapi
C 277	17	6.0	113762	9	AC005294	AC005294 Homo sapi	350	17	6.0	162607	9	AC015599	AC015599 Homo sapi
C 278	17	6.0	114388	2	AC074147	AC074147 Mus muscu	351	17	6.0	162760	9	AC007969	AC007969 Homo sapi
C 279	17	6.0	114653	9	AC092411	AC092411 Felis cat	352	17	6.0	162854	2	AL596275	AL596275 Homo sapi
C 280	17	6.0	114728	9	HSJ343124	AL133294 Human DNA	353	17	6.0	163239	2	AC091049	AC091049 Homo sapi
C 281	17	6.0	115198	9	AL139223	AL139223 Homo sapi	354	17	6.0	163338	9	AL162426	AL162426 Human DNA
C 282	17	6.0	116368	9	HS1112F19	AL034420 Human DNA	355	17	6.0	163466	9	AC010921	AC010921 Drosophill
C 283	17	6.0	117406	9	AC010792	AC010792 Homo sapi	356	17	6.0	163881	9	AC007436	AC007436 Homo sapi
C 284	17	6.0	120990	2	AC022326	AC022326 Homo sapi	357	17	6.0	164500	2	AL359698	AL359698 Homo sapi
C 285	17	6.0	124849	9	AL157414	AL157414 Human DNA	358	17	6.0	165031	2	AL445255	AL445255 Human DNA
C 286	17	6.0	125184	9	AL357141	AL357141 Human DNA	359	17	6.0	165245	9	AL451050	AL451050 Homo sapi
C 287	17	6.0	125528	2	AC027389	AC027389 Homo sapi	360	17	6.0	165909	2	AC079152	AC079152 Homo sapi
C 288	17	6.0	130965	9	AL162457	AL162457 Human DNA	361	17	6.0	165978	2	AC087132	AC087132 Mus muscu
C 289	17	6.0	135209	9	HS42616	AL020997 Human DNA	362	17	6.0	166118	9	AL355495	AL355495 Human DNA

363	17	6.0	166171	2	AL359550	Homo sapi	436	17	6.0	195382	2	AL445254	Homo sapi
364	17	6.0	166339	2	AC079186	Homo sapi	437	17	6.0	195824	2	AL590560	Homo sapi
365	17	6.0	166802	2	AL355344	Homo sapi	438	17	6.0	196286	8	ATCHRIV64	AL161564 Arabidops
366	17	6.0	167131	9	CNS01DST	Homo chr	439	17	6.0	196767	2	AC091308	Mus muscu
367	17	6.0	167515	2	AC091864	Homo sapi	440	17	6.0	196962	2	AC021161	AC021161 Homo sapi
368	17	6.0	167540	2	AC016084	Homo sapi	441	17	6.0	198470	2	AC090121	AC090121 Mus muscu
369	17	6.0	168305	2	AC024969	Homo sapi	442	17	6.0	198829	2	AL391668	AL391668 Homo sapi
370	17	6.0	170414	2	AC026019	Homo sapi	443	17	6.0	199365	2	AC068485	AC068485 Homo sapi
371	17	6.0	170545	2	AC010098	Homo sapi	444	17	6.0	199475	2	AL592182	AL592182 Homo sapi
372	17	6.0	170908	9	HSBA28009	Homo sapi	445	17	6.0	200000	2	AP000511	AP000511 Homo sapi
373	17	6.0	170984	2	AC024943	Human DNA	446	17	6.0	201175	2	AC004676	AC004676 Homo sapi
374	17	6.0	171225	2	AC092899	Homo sapi	447	17	6.0	201377	2	AC073767	AC073767 Homo sapi
375	17	6.0	171370	9	AC004021	Human PAC	448	17	6.0	201397	2	AC091628	AC091628 Homo sapi
376	17	6.0	171842	2	AC055754	Homo sapi	449	17	6.0	202309	2	AC009167	AC009167 Homo sapi
377	17	6.0	172220	2	AC021545	Homo sapi	450	17	6.0	206831	2	AC093134	AC093134 Bos tauru
378	17	6.0	172588	2	AC007430	Homo sapi	451	17	6.0	207006	2	AC055779	AC055779 Homo sapi
379	17	6.0	172825	2	AC020614	Homo sapi	452	17	6.0	208660	2	AC020577	AC020577 Homo sapi
380	17	6.0	173023	9	AL354898	Human DNA	453	17	6.0	209154	2	AL445199	AL445199 Homo sapi
381	17	6.0	173364	2	AC009807	Homo sapi	454	17	6.0	209258	2	AC019234	AC019234 Homo sapi
382	17	6.0	173501	2	AC091541	Canis fam	455	17	6.0	209669	2	AC091949	AC091949 Homo sapi
383	17	6.0	173710	2	AC024886	Homo sapi	456	17	6.0	209684	2	AL353795	AL353795 Homo sapi
384	17	6.0	173786	2	AP003069	Homo sapi	457	17	6.0	211600	2	AC013266	AC013266 Homo sapi
385	17	6.0	173971	2	AC025300	Homo sapi	458	17	6.0	212525	2	AC063977	AC063977 Homo sapi
386	17	6.0	174301	2	AP003523	Oryza sat	459	17	6.0	214001	9	AC078993	AC078993 Homo sapi
387	17	6.0	174404	2	AC013542	Homo sapi	460	17	6.0	214838	2	AC012501	AC012501 Homo sapi
388	17	6.0	174732	2	AC012160	Drosophil	461	17	6.0	217068	2	AC079434	AC079434 Mus muscu
389	17	6.0	174933	9	CNS01DRL	Human chr	462	17	6.0	219811	2	AC034123	AC034123 Mus muscu
390	17	6.0	175413	9	AC093097	Drosophil	463	17	6.0	221104	2	AC025579	AC025579 Homo sapi
391	17	6.0	175895	2	AC064862	Homo sapi	464	17	6.0	223263	2	AL158068	AL158068 Homo sapi
392	17	6.0	175958	2	AL356985	Homo sapi	465	17	6.0	224367	2	AC074336	AC074336 Mus muscu
393	17	6.0	176688	2	AC084696	Mus muscu	466	17	6.0	225300	2	AC084798	AC084798 Mus muscu
394	17	6.0	177421	2	AC032013	Mus muscu	467	17	6.0	229012	2	AC021084	AC021084 Homo sapi
395	17	6.0	177481	2	AC089999	Homo sapi	468	17	6.0	230352	2	AC016982	AC016982 Mus muscu
396	17	6.0	177736	2	AC078924	Homo sapi	469	17	6.0	235259	2	AC092275	AC092275 Homo sapi
397	17	6.0	177826	2	AC016852	Homo sapi	470	17	6.0	237674	2	AL591478	AL591478 Mus muscu
398	17	6.0	178714	33	AC022925	Ac022925 Homo sapi	471	17	6.0	241714	2	AC083834	AC083834 Mus muscu
399	17	6.0	179765	2	AC026223	Homo sapi	472	17	6.0	245383	2	AL161653	AL161653 Homo sapi
400	17	6.0	180041	2	AC009831	Homo sapi	473	17	6.0	254437	9	AP000500	AP000500 Homo sapi
401	17	6.0	180549	2	AL596087	Homo sapi	474	17	6.0	256673	2	AC087146	AC087146 Mus muscu
402	17	6.0	180848	2	AC023996	Homo sapi	475	17	6.0	258166	3	AE003652	AE003652 Drosophil
403	17	6.0	181386	9	AC007686	Homo sapi	476	17	6.0	272351	2	AL590968	AL590968 Mus muscu
404	17	6.0	181538	9	AC022161	Homo sapi	477	17	6.0	296514	3	AE003504	AE003504 Drosophil
405	17	6.0	181727	2	AC036232	Homo sapi	478	17	6.0	300739	3	AE003422	AE003422 Drosophil
406	17	6.0	182127	2	AC090137	Homo sapi	479	17	6.0	308711	2	AL358788	AL358788 Homo sapi
407	17	6.0	182250	2	CNS01DST	Homo sapi	480	17	6.0	340000	9	HS21C067	HS21C067 Homo sapi
408	17	6.0	182267	2	AC007638	Homo sapi	481	17	6.0	340000	9	AP001716	AP001716 Homo sapi
409	17	6.0	182289	2	AC021241	Homo sapi	482	17	6.0	340000	9	AP001728	AP001728 Homo sapi
410	17	6.0	183304	2	AC092625	Homo sapi	483	17	6.0	340000	9	AP001760	AP001760 Homo sapi
411	17	6.0	183428	2	AC036167	Homo sapi	484	17	6.0	348050	9	HS225042	HS225042 Homo sapi
412	17	6.0	182501	2	AC011221	Homo sapi	485	16	5.7	141	10	MMU06145	U06145 Mus muscu
413	17	6.0	182587	2	AC011439	Homo sapi	486	16	5.7	158	9	AF076940	AF076940 Homo sapi
414	17	6.0	182662	2	AC023860	Homo sapi	487	16	5.7	188	4	AF085171	AF085171 Oryctolag
415	17	6.0	184599	2	AC093259	Homo sapi	488	16	5.7	201	11	G59716	G59716 SHGC-130468
416	17	6.0	184681	2	AC084409	Mus muscu	489	16	5.7	205	6	AY059902	AY059902 Sequence
417	17	6.0	184852	2	AC026032	Homo sapi	490	16	5.7	224	8	AY022674	AY022674 Oryza sat
418	17	6.0	186092	2	AC006582	Homo sapi	491	16	5.7	225	9	HSATPCB05	AT165286 Homo sapi
419	17	6.0	186271	2	AC024168	Homo sapi	492	16	5.7	282	10	G22231	G22231 human STS W
420	17	6.0	187774	2	AL390994	Homo sapi	493	16	5.7	286	11	MUSBGAL15	M75121 Mouse acid
421	17	6.0	187999	2	AC026928	Homo sapi	494	16	5.7	288	9	HS16H9F	Z59473 H.sapiens C
422	17	6.0	188169	2	AL365394	Homo sapi	495	16	5.7	298	11	G44896	G44896 Z1168.1 zeb
423	17	6.0	188284	2	AC073344	Homo sapi	496	16	5.7	362	1	AY006676	AY006676 Unculture
424	17	6.0	188359	9	AC006636	Homo sapi	497	16	5.7	389	11	G41511	G41511 Z4373 Zebra
425	17	6.0	188544	2	AL162586	Homo sapi	498	16	5.7	390	6	AX122743	AX122743 Sequence
426	17	6.0	189181	2	AP001333	Homo sapi	499	16	5.7	430	10	MUSMGC	Z84740 H.sapiens f
427	17	6.0	189810	2	AL358177	Homo sapi	500	16	5.7	450	10	MUSMGC	D50412 H.sapiens f
428	17	6.0	190393	9	AC009102	Homo sapi	501	16	5.7	518	11	G31916	G31916 MCM242 Chic
429	17	6.0	190784	2	AC048383	Homo sapi	502	16	5.7	569	6	AR022415	AR022415 Sequence
430	17	6.0	191239	2	AP000883	Homo sapi	503	16	5.7	569	6	AR068893	AR068893 Sequence
431	17	6.0	192021	9	AP003536	Oryza sat	504	16	5.7	569	6	AR103350	AR103350 Sequence
432	17	6.0	192867	9	AL358815	Human DNA	505	16	5.7	569	6	AR134920	AR134920 Sequence
433	17	6.0	193159	2	AC006946	Homo sapi	506	16	5.7	569	6	AR141749	AR141749 Sequence
434	17	6.0	194103	9	AC060818	Homo sapi	507	16	5.7	569	6	AR143266	AR143266 Sequence
435	17	6.0	195274	9	AL445528	Human DNA	508	16	5.7	569	6	AR151875	AR151875 Sequence

C 509	16	5.7	569	6	136355	I36355 Sequence 44	582	16	5.7	2337	5	AF083382	AF083382 Danio rer
C 510	16	5.7	569	6	188055	I88055 Sequence 44	C 583	16	5.7	2350	5	CHKPRAR1A1	L11264 Gallus gall
C 511	16	5.7	650	5	AF163849	AF163849 Rata egla	C 584	16	5.7	2351	6	I64752	I64752 Sequence 1
C 512	16	5.7	658	1	HSTECTXT10	U34376 Human tyros	C 585	16	5.7	2360	10	MUSB3445	M57732 Mouse beta-
C 513	16	5.7	753	1	SCN2ARPA	X79605 S.griseus (C 586	16	5.7	2402	10	RNU02322	U02322 Rattus norv
C 514	16	5.7	784	10	MMY19226	Y19226 Mus musculu	C 587	16	5.7	2431	6	AR059883	AR059883 Sequence
C 515	16	5.7	867	9	HSU02329	U02329 Human clone	C 588	16	5.7	2431	6	I50104	I50104 Sequence 25
C 516	16	5.7	897	10	AF194440	AF194440 Rattus no	C 589	16	5.7	2449	5	AF045656	AF045656 Gallus ga
C 517	16	5.7	919	8	AF093631	AF093631 Oryza sat	C 590	16	5.7	2455	10	MAU96612	U96612 Mesocricetu
C 518	16	5.7	959	6	I79510	I79510 Sequence 1	C 591	16	5.7	2490	6	AR059881	AR059881 Sequence
C 519	16	5.7	977	10	MMU250394	AJ250394 Mus muscu	C 592	16	5.7	2490	6	I50102	I50102 Sequence 23
C 520	16	5.7	1051	9	HSU02325	U02325 Human clone	C 593	16	5.7	2501	9	HUMHERGC	M94167 Human hereg
C 521	16	5.7	1062	10	BC004726	BC004726 Mus muscu	C 594	16	5.7	2501	8	SCYAP1	X58633 Yeast YAP1
C 522	16	5.7	1069	8	CFPA233688	AJ233688 Canis fam	C 595	16	5.7	2525	10	RNU02317	U02317 Rattus norv
C 523	16	5.7	1079	8	AF050675	AF050675 Oryza sat	C 596	16	5.7	2540	10	AF194439	AF194439 Rattus no
C 524	16	5.7	1149	9	AK025378	AK025378 Homo sapi	C 597	16	5.7	2542	10	RAREFCGA	M22823 Rat Ige fc
C 525	16	5.7	1255	10	RNU02320	U02320 Rattus norv	C 598	16	5.7	2545	6	AR123798	AR123798 Sequence
C 526	16	5.7	1317	9	HUMCOL25	U76388 Human 180 k	C 599	16	5.7	2545	6	I09687	I09687 Sequence 3
C 527	16	5.7	1318	10	AF095927	AF095927 Rattus no	C 600	16	5.7	2579	8	AB025353	AB025353 Allium ma
C 528	16	5.7	1362	1	STU92525	U92525 Salmonella	C 601	16	5.7	2700	8	SCSNO3	X63268 S.cerevisia
C 529	16	5.7	1403	6	AR113041	AR113041 Sequence	C 602	16	5.7	2709	1	ECUVRD01	X00738 E.coli uvrd
C 530	16	5.7	1422	9	AY024365	AY024365 Homo sapi	C 603	16	5.7	2720	3	AY032678	AY032678 Toxoplas
C 531	16	5.7	1436	9	HSM801818	AL136850 Homo sapi	C 604	16	5.7	2726	10	CRHSP60A	M34562 C.griseus C
C 532	16	5.7	1453	9	AK001043	AK001043 Homo sapi	C 605	16	5.7	2737	10	RNU02321	U02321 Rattus norv
C 533	16	5.7	1561	14	U87662	U87662 Hepatitis G	C 606	16	5.7	2793	14	AF200624	AF200624 Peanut st
C 534	16	5.7	1611	10	AB024573	AB024573 Mus muscu	C 607	16	5.7	2805	5	AK021777	AK021777 Homo sapi
C 535	16	5.7	1637	6	AR127148	AR127148 Sequence	C 608	16	5.7	2810	4	CF35CGMP	Z68340 C.familiari
C 536	16	5.7	1637	6	AX003001	AX003001 Sequence	C 609	16	5.7	2846	1	ECUVRD02	D00069 Escherichia
C 537	16	5.7	1637	6	E36332	E36332 METPR0 02 b	C 610	16	5.7	2846	1	ECUVRD02	X04037 Escherichia
C 538	16	5.7	1640	10	MUSTFID	D01034 Mus musculu	C 611	16	5.7	2856	10	AF194438	AF194438 Rattus no
C 539	16	5.7	1750	14	BC0293276	AJ293276 Bean comm	C 612	16	5.7	2859	8	SCPRP6	X5345 Yeast prp6
C 540	16	5.7	1752	6	AX099150	AX099150 Sequence	C 613	16	5.7	2862	9	AK001491	AK001491 Homo sapi
C 541	16	5.7	1761	4	AF153198	AF153198 Canis fam	C 614	16	5.7	2871	1	AF194331	AF194331 Myxococcu
C 542	16	5.7	1765	14	AF200623	AF200623 Peanut st	C 615	16	5.7	2885	10	AF168991	AF168991 Mus muscu
C 543	16	5.7	1777	10	AF005051	AF005051 Mus muscu	C 616	16	5.7	2897	8	SCPARIG	X60780 S.cerevisia
C 544	16	5.7	1793	9	HSU02326	U02326 Human clone	C 617	16	5.7	2908	10	RNU02316	RNU02316 Rattus norv
C 545	16	5.7	1819	9	AK025360	AK025360 Homo sapi	C 618	16	5.7	2950	4	CFGCMPPA	Y13282 Canis famli
C 546	16	5.7	1852	9	HSU02328	U02328 Human clone	C 619	16	5.7	3076	9	HSCFS21	AF136274 Homo sapi
C 547	16	5.7	1856	10	MMU63933	U63933 Mus musculu	C 620	16	5.7	3105	5	AF194993	AF194993 Rattus no
C 548	16	5.7	1865	10	RNU02324	U02324 Rattus norv	C 621	16	5.7	3105	5	AF076618	AF076618 Xenopus l
C 549	16	5.7	1867	10	RATNEUDIFF	M92430 Rat neu dif	C 622	16	5.7	3136	10	RNU02323	RNU02323 Rattus norv
C 550	16	5.7	1899	6	I79511	I79511 Sequence 2	C 623	16	5.7	3187	5	AF045655	AF045655 Gallus ga
C 551	16	5.7	1907	9	HSMB01272	AL122112 Homo sapi	C 624	16	5.7	3211	5	AF045654	AF045654 Gallus ga
C 552	16	5.7	1929	9	AF062085	AF062085 Homo sapi	C 625	16	5.7	3211	10	U00678	U00678 Mus musculu
C 553	16	5.7	1945	9	AF055021	AF055021 Homo sapi	C 626	16	5.7	3216	4	CFU52868	U52868 Canis famli
C 554	16	5.7	1949	10	AF288289	AF288289 Rattus no	C 627	16	5.7	3255	9	AK000694	AK000694 Homo sapi
C 555	16	5.7	1991	6	AX155059	AX155059 Sequence	C 628	16	5.7	3272	10	RNU02318	U02318 Rattus norv
C 556	16	5.7	1995	10	BC003998	BC003998 Mus muscu	C 629	16	5.7	3507	5	AF142632	AF142632 Xenopus l
C 557	16	5.7	2000	9	BC008871	BC008871 Homo sapi	C 630	16	5.7	3568	14	AF049705	AF049705 Fijl dise
C 558	16	5.7	2010	6	AR026610	AR026610 Sequence	C 631	16	5.7	3746	8	SCYBR055C	Z35924 S.cerevisia
C 559	16	5.7	2010	6	AR029108	AR029108 Sequence	C 632	16	5.7	3761	8	SCPDRA	X53830 Yeast PDRA
C 560	16	5.7	2010	6	AR053347	AR053347 Sequence	C 633	16	5.7	3850	10	MMU41765	U41765 Mus musculu
C 561	16	5.7	2010	6	AR059878	AR059878 Sequence	C 634	16	5.7	4311	9	HSEB2F7	AL022338 Human DNA
C 562	16	5.7	2010	6	I50099	I50099 Sequence 12	C 635	16	5.7	4332	6	AR104511	AR104511 Sequence
C 563	16	5.7	2011	6	AX076679	AX076679 Sequence	C 636	16	5.7	4489	8	SCYNR069C	Z71684 S.cerevisia
C 564	16	5.7	2096	10	AB041554	AB041554 Mus muscu	C 637	16	5.7	4799	9	AB028988	AB028988 Homo sapi
C 565	16	5.7	2156	6	AB049870	AB049870 Macaca fa	C 638	16	5.7	4894	8	SCYBR056W	Z35925 S.cerevisia
C 566	16	5.7	2164	6	AR034654	AR034654 Sequence	C 639	16	5.7	5052	2	AC093061	AC093061 Homo sapi
C 567	16	5.7	2164	6	I29959	I29959 Sequence 6	C 640	16	5.7	5077	14	KSU82242	U82242 Kaposi's sa
C 568	16	5.7	2164	9	S68256	S68256 erbB2-gp30/	C 641	16	5.7	5123	3	AF093584	AF093584 Plasmodu
C 569	16	5.7	2187	9	AK022650	AK022650 Homo sapi	C 642	16	5.7	5143	6	AR035960	AR035960 Sequence
C 570	16	5.7	2196	6	AK022583	AK022583 Homo sapi	C 643	16	5.7	5143	6	AR038846	AR038846 Sequence
C 571	16	5.7	2199	6	AR034655	AR034655 Sequence	C 644	16	5.7	5143	6	HSU24170	U24170 Human p21 (
C 572	16	5.7	2199	6	AR059877	AR059877 Sequence	C 645	16	5.7	5270	10	RATTSURA	M34842 Rat thyroid
C 573	16	5.7	2199	6	I29960	I29960 Sequence 7	C 646	16	5.7	5338	9	AB051547	AB051547 Homo sapi
C 574	16	5.7	2199	6	I50098	I50098 Sequence 7	C 647	16	5.7	5481	9	AB028959	AB028959 Homo sapi
C 575	16	5.7	2199	9	HUMHERGB	M94166 Human hereg	C 648	16	5.7	5686	1	AF290950	AF290950 Pseudom
C 576	16	5.7	2224	9	HUMHERGA	M94165 Human hereg	C 649	16	5.7	5693	10	MPPLD2G4	AF052294 Mus muscu
C 577	16	5.7	2277	9	AK001712	AK001712 Homo sapi	C 650	16	5.7	5891	5	AB001981	AB001981 Columba l
C 578	16	5.7	2307	3	PEAMAI	M29898 P.fragile a	C 651	16	5.7	5982	10	D86950	D86950 Mus musculu
C 579	16	5.7	2315	9	AK022742	AK022742 Homo sapi	C 652	16	5.7	6388	10	AF24719352	AF24719352 Mus muscu
C 580	16	5.7	2318	8	AB049146	AB049146 Clavispor	C 653	16	5.7	6786	9	AB055982	AB055982 Homo sapi
C 581	16	5.7	2328	10	RNU02319	U02319 Rattus norv	C 654	16	5.7	6880	9	HSA012376	AJ012376 Homo sapi

C 655	16	5.7	6942	10	RNIGMPOR	C 728	16	5.7	38253	9	AC006276	AC006276 Homo sapi
C 656	16	5.7	6998	10	AF113751	C 729	16	5.7	38429	9	AC000035	AC000035 Homo sapi
C 657	16	5.7	7448	10	SS038915	C 730	16	5.7	38731	3	U20864	U20864 Caenorhabdi
C 658	16	5.7	7660	6	AX059884	C 731	16	5.7	39353	3	CEP11A1	CEP11A1 Caenorhabdi
C 659	16	5.7	7763	2	AC023307	C 732	16	5.7	39481	2	AC006176	AC006176 Homo sapi
C 660	16	5.7	7791	1	AE006694	C 733	16	5.7	39658	2	HS0231B4	HS0231B4 Homo sapi
C 661	16	5.7	7847	8	SCYNR070W	C 734	16	5.7	39838	2	AC007766	AC007766 Homo sapi
C 662	16	5.7	7860	6	AX092594	C 735	16	5.7	39948	2	AC011549	AC011549 Homo sapi
C 663	16	5.7	7862	6	AX135712	C 736	16	5.7	40311	2	AC074212	AC074212 Homo sapi
C 664	16	5.7	8438	1	KPN250891	C 737	16	5.7	41563	2	AC027010	AC027010 Homo sapi
C 665	16	5.7	8481	14	KSU40394	C 738	16	5.7	41860	9	AC019319	AC019319 Homo sapi
C 666	16	5.7	9015	1	U67561	C 739	16	5.7	43303	9	HS019H10	HS019H10 Homo sapi
C 667	16	5.7	9181	6	AX127771	C 740	16	5.7	43341	2	AC014471	AC014471 Drosophila
C 668	16	5.7	9181	6	AX139758	C 741	16	5.7	44368	5	AF301601	AF301601 Takifugu
C 669	16	5.7	9190	6	HSSEF1EX28	C 742	16	5.7	43731	9	AC000393	AC000393 Genomic s
C 670	16	5.7	9495	6	AX059978	C 743	16	5.7	44025	3	L15201	L15201 Caenorhabdi
C 671	16	5.7	9497	6	AX059973	C 744	16	5.7	45627	9	AL135907	AL135907 Human DNA
C 672	16	5.7	9497	6	AF165281	C 745	16	5.7	51589	9	AL135311	AL135311 Human DNA
C 673	16	5.7	9593	6	AX059976	C 746	16	5.7	52395	2	AC067783	AC067783 Homo sapi
C 674	16	5.7	9741	6	AX127830	C 747	16	5.7	53797	2	AP001415	AP001415 Homo sapi
C 675	16	5.7	9741	6	AX139817	C 748	16	5.7	55351	9	AC010154	AC010154 Homo sapi
C 676	16	5.7	9854	6	AX127831	C 749	16	5.7	58672	9	AC006524	AC006524 Homo sapi
C 677	16	5.7	9854	6	AX139818	C 750	16	5.7	58700	9	AC024939	AC024939 Homo sapi
C 678	16	5.7	9988	10	AF297869S1	C 751	16	5.7	58725	9	AC003988	AC003988 Human PAC
C 679	16	5.7	10029	1	AE005535	C 752	16	5.7	59102	9	AL133324	AL133324 Human DNA
C 680	16	5.7	10469	1	AE000457	C 753	16	5.7	60117	2	AC016244	AC016244 Homo sapi
C 681	16	5.7	10551	1	AE000458	C 754	16	5.7	61417	2	AC023564	AC023564 Homo sapi
C 682	16	5.7	10602	1	AE004562	C 755	16	5.7	61676	2	AC060800	AC060800 Homo sapi
C 683	16	5.7	10961	1	AE001939	C 756	16	5.7	62281	2	AC010304	AC010304 Homo sapi
C 684	16	5.7	11106	10	MMSAP62	C 757	16	5.7	62470	2	AC015828	AC015828 Homo sapi
C 685	16	5.7	11307	1	AE000387	C 758	16	5.7	63595	2	AC084774	AC084774 Homo sapi
C 686	16	5.7	11392	1	AE003940	C 759	16	5.7	65792	6	AX067456	AX067456 Sequence
C 687	16	5.7	11505	1	AE005940	C 760	16	5.7	66440	2	AC027791	AC027791 Homo sapi
C 688	16	5.7	11940	14	MYNPMGLCE	C 761	16	5.7	66623	2	AC079057	AC079057 Homo sapi
C 689	16	5.7	12016	2	AC008789	C 762	16	5.7	67332	9	AP001468	AP001468 Homo sapi
C 690	16	5.7	12184	2	AC020793	C 763	16	5.7	67847	2	AC017241	AC017241 Drosophila
C 691	16	5.7	13182	1	AE005790	C 764	16	5.7	68248	2	AC090295	AC090295 Homo sapi
C 692	16	5.7	13646	9	HUMRBP1	C 765	16	5.7	68248	2	AC090295	AC090295 Homo sapi
C 693	16	5.7	13742	10	MMCMW17	C 766	16	5.7	68484	8	NCB17C10	NCB17C10 Homo sapi
C 694	16	5.7	13865	2	AC014472	C 767	16	5.7	68848	1	MTV043	MTV043 Mycobacte
C 695	16	5.7	14296	1	AF097520	C 768	16	5.7	68973	2	AC090920	AC090920 Homo sapi
C 696	16	5.7	14951	1	AE000756	C 769	16	5.7	69012	2	AC023849	AC023849 Homo sapi
C 697	16	5.7	15382	1	AE006974	C 770	16	5.7	69012	2	AC023849	AC023849 Homo sapi
C 698	16	5.7	18376	9	AL354800	C 771	16	5.7	69528	2	AF146367	AF146367 Homo sapi
C 699	16	5.7	19798	9	AB014460	C 772	16	5.7	69992	2	AC067905	AC067905 Homo sapi
C 700	16	5.7	20083	9	HS01961	C 773	16	5.7	70326	2	AC068957	AC068957 Homo sapi
C 701	16	5.7	20556	9	HS01961	C 774	16	5.7	70422	2	AP001961	AP001961 Homo sapi
C 702	16	5.7	23652	2	AP004034	C 775	16	5.7	71230	9	AC004508	AC004508 Homo sapi
C 703	16	5.7	23771	1	AF141323	C 776	16	5.7	72485	2	AC083792	AC083792 Homo sapi
C 704	16	5.7	23920	8	SCA10131	C 777	16	5.7	72511	2	AC025954	AC025954 Homo sapi
C 705	16	5.7	25638	1	ECU28379	C 778	16	5.7	72533	2	AC013619	AC013619 Homo sapi
C 706	16	5.7	26700	6	AR068044	C 779	16	5.7	72842	2	AC083957	AC083957 Homo sapi
C 707	16	5.7	26700	6	AR068044	C 780	16	5.7	73431	2	AP000575	AP000575 Homo sapi
C 708	16	5.7	26802	2	AC018087	C 781	16	5.7	73511	9	HS65A6	HS65A6 Human DNA
C 709	16	5.7	28026	9	AC079743	C 782	16	5.7	73994	2	AC079000	AC079000 Homo sapi
C 710	16	5.7	28244	9	AC002481	C 783	16	5.7	74224	2	AC021724	AC021724 Homo sapi
C 711	16	5.7	28559	14	AF148805	C 784	16	5.7	74307	2	AL353760	AL353760 Mus muscu
C 712	16	5.7	29366	1	SC9571X	C 785	16	5.7	74332	2	AC073764	AC073764 Mus muscu
C 713	16	5.7	29372	2	AC018117	C 786	16	5.7	74990	2	AC024421	AC024421 Homo sapi
C 714	16	5.7	31495	1	BU033883	C 787	16	5.7	75436	2	AC010678	AC010678 Homo sapi
C 715	16	5.7	31570	9	AL356424	C 788	16	5.7	75475	8	AB005235	AB005235 Arabidops
C 716	16	5.7	32017	2	AC013048	C 789	16	5.7	75546	2	AP000460	AP000460 Homo sapi
C 717	16	5.7	32207	6	AR065852	C 790	16	5.7	75689	2	AC018536	AC018536 Homo sapi
C 718	16	5.7	32207	6	AR065852	C 791	16	5.7	76795	2	AC023515	AC023515 Homo sapi
C 719	16	5.7	32923	1	SCD82	C 792	16	5.7	76829	2	AC027007	AC027007 Homo sapi
C 720	16	5.7	33026	2	AC008988	C 793	16	5.7	77902	2	AC006958	AC006958 Homo sapi
C 721	16	5.7	33934	10	MMSINDEXA	C 794	16	5.7	78843	2	AC015455	AC015455 Homo sapi
C 722	16	5.7	34713	1	SCF42	C 795	16	5.7	79489	2	AC010410	AC010410 Homo sapi
C 723	16	5.7	35213	9	HS9B17	C 796	16	5.7	79562	2	AC005600	AC005600 Homo sapi
C 724	16	5.7	36154	5	FRU301641	C 797	16	5.7	79622	2	AC036119	AC036119 Homo sapi
C 725	16	5.7	36612	2	AL354764	C 798	16	5.7	79951	2	HSU71B4	HSU71B4 Human DNA
C 726	16	5.7	37077	9	AF135116	C 799	16	5.7	80068	2	AC022435	AC022435 Homo sapi
C 727	16	5.7	37223	9	AF135116	C 800	16	5.7	80674	9	AL355599	AL355599 Homo sapi

C 801	16	5.7	81854	9	AC006396	AC006396 Homo sapi	C 874	16	5.7	107808	2	AL162271	AL162271 Homo sapi
C 802	16	5.7	82091	2	AC024592	AC024592 Homo sapi	C 875	16	5.7	108369	17	AF124523	AF124523 Homo sapi
C 803	16	5.7	82956	2	AC068987	AC068987 Homo sapi	C 876	16	5.7	109173	2	AC007678	AC007678 Homo sapi
C 804	16	5.7	83549	9	AC004752	AC004752 Homo sapi	C 877	16	5.7	109290	2	HS838L14	Y12335 Homo sapien
C 805	16	5.7	83820	9	AL133264	AL133264 Homo sapi	C 878	16	5.7	109381	9	AP002350	AP002350 Homo sapi
C 806	16	5.7	84684	9	AC035142	AC035142 Homo sapi	C 879	16	5.7	109880	2	AF181895	AF181895 Homo sapi
C 807	16	5.7	84873	2	AC024570	AC024570 Homo sapi	C 880	16	5.7	110000	2	AC008576_1	Continuation (2 of
C 808	16	5.7	85217	8	AF117J13	AL138651 Arabidops	C 881	16	5.7	110000	2	AC046137_0	AC046137 Homo sapi
C 809	16	5.7	86372	8	AC003113	AC003113 Genomic s	C 882	16	5.7	110000	2	AC046137_0	AC046137 Homo sapi
C 810	16	5.7	86437	2	AC090276	AC090276 Homo sapi	C 883	16	5.7	110000	2	AC009801_0	AC009801 Homo sapi
C 811	16	5.7	86514	9	AL136117	AL136117 Homo sapi	C 884	16	5.7	110000	2	AL512625_2	Continuation (3 of
C 812	16	5.7	86535	9	HS491M17	Z97988 Human DNA S	C 885	16	5.7	110000	2	LMFLCHR36_00	AL499624 Leisman1
C 813	16	5.7	86719	9	HSJ906P16	AL079339 Human DNA	C 886	16	5.7	110000	2	LMFLCHR36_01	Continuation (2 of
C 814	16	5.7	87102	2	AF205591	AF205591 Homo sapi	C 887	16	5.7	110000	2	LMFLCHR36_16	Continuation (17 o
C 815	16	5.7	87581	8	ATT7H20	AL165508 Arabidops	C 888	16	5.7	110000	2	LMFLCHR36_19	Continuation (20 o
C 816	16	5.7	87925	9	AC005318	AC005318 Homo sapi	C 889	16	5.7	110000	2	AC022133_4	Continuation (5 of
C 817	16	5.7	87942	9	HS310J6	AL035593 Human DNA	C 890	16	5.7	111345	9	AL445531	AL445531 Human DNA
C 818	16	5.7	88046	9	AP000965	AP000965 Homo sapi	C 891	16	5.7	111678	2	AC026209	AC026209 Homo sapi
C 819	16	5.7	89319	9	HSBK29P11	AL118516 Human DNA	C 892	16	5.7	112027	2	AC007006	AC007006 Homo sapi
C 820	16	5.7	90429	9	HS55110	AL035408 Human DNA	C 893	16	5.7	112080	9	HS482L3	AL031446 Human DNA
C 821	16	5.7	91387	8	ATF7J7	AL021960 Arabidops	C 894	16	5.7	112424	2	AC022502	AC022502 Homo sapi
C 822	16	5.7	91414	1	ECOWH85	M87045 E. coli gen	C 895	16	5.7	112424	2	AC022502	AC022502 Homo sapi
C 823	16	5.7	92882	9	AL355815	AL355815 Human DNA	C 896	16	5.7	112595	9	AF196970	AF196970 Homo sapi
C 824	16	5.7	92884	2	AC092140	AC092140 Homo sapi	C 897	16	5.7	112621	9	AC004891	AC004891 Homo sapi
C 825	16	5.7	93581	9	HS1059H15	AL022100 Human DNA	C 898	16	5.7	112910	2	AL590235	AL590235 Homo sapi
C 826	16	5.7	94153	9	AL590037	AL590037 Human DNA	C 899	16	5.7	112961	9	HS322112	HS322112 Homo sapi
C 827	16	5.7	94376	9	AC005827	AC005827 Homo sapi	C 900	16	5.7	113168	9	HSJ779E11	HSJ779E11 Homo sapi
C 828	16	5.7	94409	2	AC022305	AC022305 Homo sapi	C 901	16	5.7	113644	9	HSJD129P5	HSJD129P5 Homo sapi
C 829	16	5.7	94459	10	MMT5XDNA	X95946 M. musculus	C 902	16	5.7	114314	9	AL359494	AL359494 Human DNA
C 830	16	5.7	95484	2	AC015854	AC015854 Homo sapi	C 903	16	5.7	114361	2	AC034232	AC034232 Homo sapi
C 831	16	5.7	95801	9	AC078776	AC078776 Homo sapi	C 904	16	5.7	114774	2	AL603748	AL603748 Homo sapi
C 832	16	5.7	96086	1	STYSTMD1	AF233324 Salimone11	C 905	16	5.7	115039	9	HS145B8	AL035451 Human DNA
C 833	16	5.7	96360	2	AP003867	AP003867 Oryza sat	C 906	16	5.7	115295	9	AL136967	AL136967 Homo sapi
C 834	16	5.7	96381	2	AC068610	AC068610 Homo sapi	C 907	16	5.7	115583	9	HSJ351K20	AL109939 Human DNA
C 835	16	5.7	96400	9	AL356516	AL356516 Human DNA	C 908	16	5.7	115857	9	HS816K17	AL031678 Human DNA
C 836	16	5.7	96678	9	HS477J10A	AL0196755 Human DNA	C 909	16	5.7	115915	9	AP001464	AP001464 Homo sapi
C 837	16	5.7	96914	2	AC010247	AC010247 Homo sapi	C 910	16	5.7	116069	2	AP000784	AP000784 Homo sapi
C 838	16	5.7	96936	8	ATF8R6	AL162873 Arabidops	C 911	16	5.7	117431	9	HS850H21	AL031680 Human DNA
C 839	16	5.7	97112	9	AL353631	AL353631 Human DNA	C 912	16	5.7	117655	2	AC0744221	AC0744221 Mus muscu
C 840	16	5.7	97687	9	AC008642	AC008642 Homo sapi	C 913	16	5.7	118153	9	HS249C1	AL022154 Human DNA
C 841	16	5.7	98246	9	AL353699	AL353699 Human DNA	C 914	16	5.7	118511	34	AC009747	AC009747 Drosophill
C 842	16	5.7	98606	2	AC015916	AC015916 Homo sapi	C 915	16	5.7	118572	9	AL133408	AL133408 Human DNA
C 843	16	5.7	98908	9	AL139279	AL139279 Human DNA	C 916	16	5.7	118802	9	AC005481	AC005481 Homo sapi
C 844	16	5.7	99163	9	AL355602	AL355602 Human DNA	C 917	16	5.7	119656	2	AP003772	AP003772 Homo sapi
C 845	16	5.7	99336	2	AC007411	AC007411 Homo sapi	C 918	16	5.7	120007	9	AF064864	AF064864 Homo sapi
C 846	16	5.7	99558	2	AP000668	AP000668 Homo sapi	C 919	16	5.7	120385	9	HS1437116	AL049570 Human DNA
C 847	16	5.7	99669	2	AC093205	AC093205 Homo sapi	C 920	16	5.7	120387	9	AC003051	AC003051 Homo sapi
C 848	16	5.7	100000	9	AB020861	AB020861 Homo sapi	C 921	16	5.7	120609	9	AF165176	AF165176 Homo sapi
C 849	16	5.7	100000	9	AB020866	AB020866 Homo sapi	C 922	16	5.7	120988	9	AL137182	AL137182 Human DNA
C 850	16	5.7	100000	9	AP000018	AP000018 Homo sapi	C 923	16	5.7	121424	9	HS418A9	Z84480 Homo sapien
C 851	16	5.7	100000	9	AP000068	AP000068 Homo sapi	C 924	16	5.7	121426	9	AL391123	AL391123 Human DNA
C 852	16	5.7	100000	9	AP000159	AP000159 Homo sapi	C 925	16	5.7	121687	9	AP000478	AP000478 Homo sapi
C 853	16	5.7	100259	9	AC005800	AC005800 Homo sapi	C 926	16	5.7	121774	2	AC084059	AC084059 Mus muscu
C 854	16	5.7	100578	9	AL390057	AL390057 Human DNA	C 927	16	5.7	121902	9	AC008386	AC008386 Homo sapi
C 855	16	5.7	101248	2	AP003884	AP003884 Oryza sat	C 928	16	5.7	122113	2	AC007573	AC007573 Drosophill
C 856	16	5.7	101270	2	HS355C18	AL023327 Human DNA	C 929	16	5.7	122134	2	AP000819	AP000819 Homo sapi
C 857	16	5.7	102144	2	AC068316	AC068316 Homo sapi	C 930	16	5.7	122364	9	AC007324	AC007324 Homo sapi
C 858	16	5.7	102818	2	AP000621	AP000621 Homo sapi	C 931	16	5.7	123004	9	AC073846	AC073846 Homo sapi
C 859	16	5.7	103181	9	HS97P20	AL031297 Human DNA	C 932	16	5.7	123004	9	AC073846	AC073846 Homo sapi
C 860	16	5.7	103268	9	HS67M12	AL008732 Human DNA	C 933	16	5.7	123320	9	AP000863	AP000863 Homo sapi
C 861	16	5.7	103699	9	AC034305	AC034305 Homo sapi	C 934	16	5.7	123782	9	HSJD756N5	HSJD756N5 Homo sapi
C 862	16	5.7	104339	2	AC018620	AC018620 Homo sapi	C 935	16	5.7	123828	9	AP000353	AP000353 Homo sapi
C 863	16	5.7	104641	9	AC026448	AC026448 Homo sapi	C 936	16	5.7	124000	2	AC004060	AC004060 Homo sapi
C 864	16	5.7	104749	2	AC091820	AC091820 Homo sapi	C 937	16	5.7	124347	2	AC010072	AC010072 Homo sapi
C 865	16	5.7	105872	2	AC091993	AC091993 Homo sapi	C 938	16	5.7	124456	2	AC017830	AC017830 Drosophill
C 866	16	5.7	106265	9	AL133294	AL133294 Human DNA	C 939	16	5.7	124506	2	AC090040	AC090040 Homo sapi
C 867	16	5.7	106290	2	AL513477	AL513477 Homo sapi	C 940	16	5.7	124510	9	AC027306	AC027306 Homo sapi
C 868	16	5.7	106300	9	AC008602	AC008602 Homo sapi	C 941	16	5.7	124940	9	HS162E17	AL160492 Homo sapi
C 869	16	5.7	106350	2	AL139413	AL139413 Homo sapi	C 942	16	5.7	125331	2	AC008820	AC008820 Homo sapi
C 870	16	5.7	106604	9	HSXKSRPX	AL239329 Homo sapi	C 943	16	5.7	125685	9	AC073310	AC073310 Homo sapi
C 871	16	5.7	106687	9	AL356124	AL356124 Human DNA	C 944	16	5.7	125850	2	AC010003	AC010003 Drosophill
C 872	16	5.7	107191	2	AC034221	AC034221 Homo sapi	C 945	16	5.7	125977	2	CNS05TC5	AL355053 Human chr
C 873	16	5.7	107808	2	AL162271	AL162271 Homo sapi	C 946	16	5.7	126000	9	AP000744	AP000744 Homo sapi

ALIGNMENTS

RESULT	1			
AX076669				
LOCUS	AX076669	2087 bp	DNA	PAT
DEFINITION	Sequence 3 from Patent WO0104145.			06-FEB-2001
ACCESSION	AX076669			
VERSION	AX076669.1	GI:12711200		
KEYWORDS	.			
SOURCE	Rattus sp.			
ORGANISM	Rattus sp.			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;			
	Rattus.			
REFERENCE	1 (bases 1 to 2087)			
AUTHORS	Thorens, B., Ibberson, M. and Oldry, M.			

FEATURES	source	Location/Qualifiers
JOURNAL	Patent: WO 0104145-A 3 18-JAN-2001; University of Lausanne (CH)	
FEATURES	source	Location/Qualifiers
CDS		/organism="Rattus sp." /db_xref="taxon:10118" 30..1466 /note="unnamed protein product" /codon_start=1 /protein_id="CAC28496.1" /db_xref="GI:12711201" /translation="MSPEDPQETQPLRSPGARAPGGRVFLATFAALGPILSFGPAL GYSSPAPLSLRRTAPPALRLDGTAAAMSGAVTGAAGVGLDGRAGKLSLIL CTVEVTFANVTAAADVMVLLGRLLTGLACGVASLVAAPYISLTAAPVARGILSGC VOLMVTGILATYVAGVYLMERKRLIADLCVPTLMLLMCMETPRRLTQHOQE MAALRFLTMSGEQWEPVGAHEGFOIAMLRRGVNHPILLIGICLAFQOLSEVNA MFVNNTIFEAKFKDSSLSAVTVGILQVLPFAVALIDRRGRRLTLALSGIVFSSNA SAFPTEKLTQSGPSSSHVGLLWPIISAEPADVHLGLAMLVGSMCLFIAGFAVGMGFI LPIMLMSEIPLRIHKGVAIVGVLTNNMFALVTKEFNSIMEILRPYCAFWLTAAFC LSVPLTLPVETKRGRTLEQITTAHEGR"
BASE COUNT	355 a 673 c 580 g 479 t	
ORIGIN		
Query Match	100.0%; Score 282; DB 6; Length 2087;	
Best Local Similarity	100.0%; Pred. No. 1.9e-146;	
Matches	282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1	gagctgctgatgttcacctggggcgtgcgcgtgcgtgcgtgtaggaagatgctcttc 60
Db	1116	GAGCCTGCTGATGATGCACCTGGGGCTGGCCCTGGCTGTAGGACGATGTGCTCTTC 1175
OY	61	atcgcctgcttttcgaatgagctcgggagaccatccctgctccatgctcaagatcttc 120
Db	1176	ATCGCTGCTTTTTCAGTAGTGCGTGGGACCCATCCCTGGCTCTCATGCACAGATCTTC 1235
OY	121	ccctcgacatcaagggtgctgagctacggcgctctgttctctccacaactggttcagcc 180
Db	1236	CCTGTGCACATCAAGAGGTGTGCTACCGCGCTCTGTCTCTCCACCAACGTGTTCATGGCC 1295
OY	181	tttcgtgtagcaaaagaattcaagacatcatgagagatccctcaagccctcagcgcttc 240
Db	1296	TTTCTGTGTCACCAAGAAGTTTAAACAGCATCATGGAGATCTTAGACCTTAGCGGCTTTC 1355
OY	241	tggctaccgcgtgcctctctgtaacctcaagcgctccctttcaag 282
Db	1356	TGGCTACCGCGCTGCTCTGTATCCTTCAGCGCTCTTTTCACG 1397
RESULT	2	
RNO245935		
LOCUS	RNO245935	2087 bp mRNA ROD 18-FEB-2000
DEFINITION	Rattus norvegicus mRNA for glucose transporter (GLUTX1 gene).	
ACCESSION	AJ245935	
VERSION	AJ245935.1	GI:7018604
KEYWORDS	glucose transporter; GLUTX1 gene.	
SOURCE	Norway rat.	
ORGANISM	Rattus norvegicus	
REFERENCE	1 (bases 1 to 2087)	
AUTHORS	Ibberson,M., Uldry,M. and Thorens,B.	
TITLE	GLUTX1, a novel mammalian glucose transporter expressed in the	
JOURNAL	central nervous system and insulin-sensitive tissues	
MEDLINE	J. Biol. Chem. 275 (7), 4607-4612 (2000)	
REFERENCE	2 (bases 1 to 2087)	
AUTHORS	Ibberson,M.R.	
TITLE	Direct Submission	
JOURNAL	Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology	
REFERENCE	and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne.	

FEATURES	CH-1005, SWITZERLAND
Source	Location/Qualifiers
gene	1..2087
CDS	/organism="Rattus norvegicus" /db_xref="taxon:10116" 30..1466 /gene="GLUTX1" 30..1466 /gene="GLUTX1" /codon_start=1 /product="glucose transporter" /protein_id="CAB5729.1" /db_xref="GI:7018605" /translation="MSPEDPEQEPOLLRSPGARAGGRVFLATFAALGPIPLSPFAL GYSPPAPSLRRTAPALRLQDTSAPGVAATGACVYLGWLLDRGRKLSILL CTVFPVTFAYITTAARDYVWMLGRLTLTGACGVAASIVAPYISEIATPARGLLGSC VQLVWVTGILLATVAGVNLKRWRLAVLGCVPPTIMLLMCKMPETTPRFLLTHOYQDEA MAALRFPMGSEEGEPPVGAEHQGFOLAMLRPGVHRPILIGILWFOOLSSVNAIM MFVANTTFFEEKFKDSSLASTVGI IQVLTFAVVALIMDRAGRKLLALSGVIMVPSM SAFCTYFKLTOSGNSHGGLVPIISAEPPDHLGLAMLAVGSMCLFAGFAVGMP IPLMLSEITPLIHKGVAATGCYVLTNMFMAFLVYKENSIMELRPYGAFLTNAFCI LSVLEFLTFVPEETKCRDTLEQTAAEEGR"
BASE COUNT	355 a 673 c 580 g 479 t
ORIGIN	

Query Match	100.0%	Score 282:	DB 10:	Length 2087:
Best Local Similarity	100.0%	Prd. No. 1.9e-146:		
Matches 282:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
Qy 1	gagctctgtaibgtacccctggagctgagccctgctgctgtaagcagcatgtctcttc 60			
Db 1116	GAGCCTGCTGATGTTCACCTGGGCGCTGGCCGGCGGCTGTAGCGAGCATGTCCTTTT 1175			
Qy 61	atcgcttgcttttgcagtagagctgggagcccatccctggctccctcatibcagaatctc 120			
Db 1176	ATCGCTGGTTTTGGCGATGAGCGTGGGAGCCATCCCTGAGCTCTCATCTCATGAGATCTTC 1235			
Qy 121	cctctgcacatcaagggtgtgtgctaacggcgatctgtctcccaaccaatlgctcatggcc 180			
Db 1236	CCTCGACATCAACAGGTGTGGCTACCGCGGCTGTCTCTCACCACACTGGTTATGGCC 1295			
Qy 181	ttctctgtgtgaccaaagaattaaacagatcatalgagatccctcagaccctgaagcgcttc 240			
Db 1296	TTTCTGGTGAGACCAAGATTTAAACAGCATCATGTGAGATCTCAGACCCCTACGGCCCTTC 1355			
Qy 241	tgcctcacgcgctgcctctctgataccctgaagcgtccctttcacg 282			
Db 1356	TGGCTACACGCTGCCTTCTGTATCTCTCAGCGGTCTTTTCACC 1397			
RESULT 3				
AB033418				
LOCUS	AB033418 2189 bp mRNA	ROD	11-APR-2000	
DEFINITION	Rattus norvegicus glurb mRNA for glucose transporter 8, complete cds.			
ACCESSION	AB033418			
VERSION	AB033418.1 GI:7592743			
KEYWORDS	glucose transporter 8.			
SOURCE	Rattus norvegicus cDNA to mRNA, clone_1lb:tests.			
ORGANISM	Rattus norvegicus			
	Eumariota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:			
	Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae:			
	Rattus.			
REFERENCE	1 (bases 1 to 2189)			
AUTHORS	Ishibashi,K.			
TITLE	Molecular cloning of a new putative glucose transporter			
JOURNAL	Published Only in Database (2000) In press			
REFERENCE	2 (bases 1 to 2189)			
AUTHORS	Ishibashi,K.			
TITLE	Direct Submission			
JOURNAL	Submitted (08-OCT-1999) to the DDBJ/EMBL/Genbank databases. Kenichi			
	Ishibashi, Jichi Medical School, Pharmacology; Miunai-Kawachi,			

FEATURES

source

gene

CDS

392 a 702 c 609 g 486 t

Kawachi, Toshiyuki 329-0498, Japan (E-mail: kashiba@ichi.ac.jp,
Tel:81-265-58-7326, Fax:81-265-44-5541)

Location/Qualifiers

1. .2189

/organism="Rattus norvegicus"

/db_xref="taxon:10116"

/clone_lib="testis"

109..1548

/gene="glut8"

109..1548

/gene="glut8"

/note="Glut8 has 12 transmembrane segments"

/codon_start=1

/product="glucose transporter 8"

/protein_id="BA94383.1"

/db_xref="GI:7592744"

/translation="MSPEDPPQETQPLLRSPGARPGRRVFLATFAALGPIILFEGFAL
GYSPALSLRRATPALRLIGDTNAPSWFAGAVITLGAAGCAVGLMGLIDBAGKRLSL
LCTPYEVTFGRVITPAARDYMLLGGRIITLACGSAVLAAPYISLTAIPYAVGILGS
CVQALMVTGILLATVACSVLLEMRILAVLGCVPPLMLLMCGIMPEITRFRLILTHOYE
AMALRLFMSEEGEEMRPVGAEHQGLQMLMRPGVHKPLIGICLIMVQOLSGVNA
IMEFANTIFPEAKFKDSSLASVTVGIQVLTVAVALIMDRAGKRLILASGIVFMS
MSAQEFYFKLQSGPSNNSHGLVPLISAEPADYHGLDLAMVAGSMCLPIACGAVGCG
PIPLMLSEITPLHITKGVATGVCLTMMFAPLVTHKEFNSIMETILRPYGAEMTLTAAC
ILSLPFLITVEPTKGRITLEDQITPAHFGR"

Query Match	Similarity	100.0%	Score	282	DB	10	Length	2189	
Best Local	Similarity	100.0%	Pred.	No. 1	9e-146				
Matches	282	Conservative	0	Mismatches	0	Indels	0	Gaps	0

QY	1	gagcgtcgtatgtatctaccctggagcctgagcctgctgagctgtatgaagcagatgtccctctc	60
Db	1198	GAGCCTGCTGATGTTTCACCTGGGCGCTGGCCCTGCGCTGTAGGCAAGCATGTGCTTTC	1257
QY	61	atcgctgtgttttgcagatgagctgtaggaccacatcccttgctccctcatgtacagatctc	120
Db	1258	ATCGCTGTGTTTTCGAGTAGGCTGGGGACCCATCCCTTGCTCTCATGTACAGATCTTC	1317
QY	121	ccctgtcacatcaagatgtgtgctacccggcctgtgtcctccacccactgtgtlcatagcc	180
Db	1318	CCTGTGCATCATCAAGGTGTGGCTACCGGCGCTGTGTCTTCACCAACTGCTTCATGGCC	1377
QY	181	tttcgtgtgaccaaagatlttaacagcatcttgagatccctcaagcccttaagggccttc	240
Db	1378	TTTCTGTGTGACCAAGATCTTAAACAGCATCATGAGATCTCTAGACCTTACGGCGCCTTC	1437
QY	241	tgagctcacgcctgcctctctgataccctcagcgcctcttlltcaag	282
Db	1438	TGGCTCACCGCTGCTTCTGTATCCTCAGCGCTTTTCACG	1479

RESULT	4		
MMU017802			
LOCUS	MMU017802	1490 bp	mRNA
DEFINITION	Mus musculus mRNA for glucose transporter 8 (GLUT8 gene).		
ACCESSION	Y17802		
VERSION	Y17802.1	GI:7688219	
KEYWORDS	glucose transporter 8; GLUT8 gene.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
	Euarchonta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 1490)		
AUTHORS	Doege, H., Schürmann, A., Bahrenberg, G., Brauers, A. and Joost, H. G.		
TITLE	GLUT8, a novel member of the sugar transport facilitator family with glucose transport activity		
JOURNAL	J. Biol. Chem.	275 (21),	16275-16280 (2000)
MEDLINE	20283667		
REFERENCE	2 (bases 1 to 1490)		
AUTHORS	Joost, H.G.		

TITLE Direct Submission
JOURNAL Submitted (16-JUL-1998) H. Joost, Institute of Pharmacology and Toxicology, Technical University Aachen, Wendlingweg 2, D-52057 Aachen, FRG

FEATURES
Source Location/Qualifiers

1. 1490
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue-type="testis"
26. 1459
/gene="Glut8"
26. 1459
/gene="Glut8"
/codon_start=1
/product="glucose transporter 8"
/protein_id="CAB89815.1"
/db_xref="GI:7688220"

gene

CDS

BASE COUNT 222 a 499 c 438 g 331 t
ORIGIN

Query Match 17.7%; Score 50; DB 10; Length 1490;
Best Local Similarity 100.0%; Pred. No. 8.7e-17;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 79 ggcctggggaaccatccctgctcctcatgtcagagatcttcctctgca 128
|||||
Db 1187 ggcctggggaaccatccctgctcctcatgtcagagatcttcctctgca 1236

RESULT 5
AF232061 1843 bp mRNA ROD 23-JUN-2000
DEFINITION Mus musculus glucose transporter GLUT8 mRNA, complete cds.
ACCESSION AF232061
VERSION AF232061.1 GI:8671757
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1843)
Carayannopoulos, M.O., Chi, M.M., Cui, Y., Pingsterhaus, J.M.,
McKnight, R.A., Mueckler, M., Devaskar, S.U. and Moley, K.H.,
GLUT8 is a glucose transporter responsible for Insulin-stimulated
glucose uptake in the blastocyst
Proc. Natl. Acad. Sci. U.S.A. 97 (13), 7313-7318 (2000)
2 (bases 1 to 1843)
Moley, K.H., Carayannopoulos, M.O. and Cui, Y.
Direct Submission
Submitted (07-FEB-2000) OB/Gyn, Washington University, 4566 Scott
Ave, St. Louis, MO 63110, USA
location/Qualifiers
1. 1843
/organism="Mus musculus"
/strain="129"
/db_xref="taxon:10090"
1. 148
49. 1482
/codon_start=1
/product="glucose transporter GLUT8"
/protein_id="AA78366.1"
/db_xref="GI:8671758"

5'UTR
CDS

/translation="MSPEDPQETQPLLRPPARTPRGRVFLASFAALGPLSGFAL
GYSSPAIPSLKRTAPALRLGDNAASWFGAVVTGAAGGILGSLDLDRGRLSLLL
CTVPVTFGFAVITARDVMMLLGRLLTGLACGVAASLAPYIIEIAPAVRGLGSC
VOLMWVIGILLAYVAGVWLEMRMLAVGCVPTMLLMCMPTPRELLTHOYOEA
MAALRFLMGSEEGEPEPVGAENHGFOLLRPGIYKPLIGISLWFOOLSGVNAI
MFANSTFEKAKFKDSSLASVTVGIIIVLFTAVALLMDRAGRLLALSGVTVFSM
SAFGTYRKLTLQSLPSNSSHVLPIAEPPDVQGLAMLVGSCFLIAGFAVGMGPI
PWLIMSEIFPLHVGAVATGICVLTNMFMALVTKFEFSSVMEMLRPGAFWLTAACAL
SVLFTLVVPEPKRRTLEQVTAHEGR"
1483. 1843
3'UTR
BASE COUNT 291 a 609 c 529 g 410 t 4 others
ORIGIN

Query Match 17.7%; Score 50; DB 10; Length 1843;
Best Local Similarity 100.0%; Pred. No. 8.6e-17;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 79 ggcctggggaaccatccctgctcctcatgtcagagatcttcctctgca 128
|||||
Db 1210 ggcctggggaaccatccctgctcctcatgtcagagatcttcctctgca 1259

RESULT 6
AX076671 2072 bp DNA PAT 06-FEB-2001
LOCUS AX076671
DEFINITION Sequence 5 from Patent WO0104145.
ACCESSION AX076671
VERSION AX076671.1 GI:12711202
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2072)
Thorens, B., Idberson, M. and Uldry, M.
Glutx polypeptide family and nucleic acids encoding same
Patent: WO 0104145-A 5 18-JAN-2001;
University of Lausanne (CH)
location/Qualifiers
1. 2072
/organism="Mus sp."
/db_xref="taxon:10095"
21. 1454
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28497.1"
/db_xref="GI:12711203"

CDS

/translation="MSPEDPQETQPLLRPPARTPRGRVFLASFAALGPLNFGFAL
GYSSPAIPSLKRTAPALRLGDNAASWFGAVVTGAAGGILGSLDLDRGRLSLLL
CTVPVTFGFAVITARDVMMLLGRLLTGLACGVAASLAPYIIEIAPAVRGLGSC
VOLMWVIGILLAYVAGVWLEMRMLAVGCVPTMLLMCMPTPRELLTHOYOEA
MAALRFLMGSEEGEPEPVGAENHGFOLLRPGIYKPLIGISLWFOOLSGVNAI
MFANSTFEKAKFKDSSLASVTVGIIIVLFTAVALLMDRAGRLLALSGVTVFSM
SAFGTYRKLTLQSLPSNSSHVLPIAEPPDVQGLAMLVGSCFLIAGFAVGMGPI
PWLIMSEIFPLHVGAVATGICVLTNMFMALVTKFEFSSVMEMLRPGAFWLTAACAL
SVLFTLVVPEPKRRTLEQVTAHEGR"
BASE COUNT 351 a 673 c 584 g 464 t
ORIGIN

Query Match 17.7%; Score 50; DB 6; Length 2072;
Best Local Similarity 100.0%; Pred. No. 8.5e-17;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 79 ggcctggggaaccatccctgctcctcatgtcagagatcttcctctgca 128
|||||
Db 1182 ggcctggggaaccatccctgctcctcatgtcagagatcttcctctgca 1231

RESULT 7
MM0245936 2072 bp mRNA ROD 18-FEB-2000
LOCUS MM0245936

DEFINITION Mus musculus mRNA for glucose transporter (GLUTX1 gene).
ACCESSION AJ245936
VERSION AJ245936.1 GI:7018606
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 2072)
TITLE GLUTX1, a novel mammalian glucose transporter expressed in the central nervous system and insulin-sensitive tissues
JOURNAL J. Biol. Chem. 275 (7), 4607-4612 (2000)
MEDLINE 20138191
REFERENCE 2 (bases 1 to 2072)
AUTHORS Ibberson, M.R.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne, CH-1005, SWITZERLAND
FEATURES
source Location/Qualifiers
1..2072
/organism="Mus musculus"
/db_xref="taxon:10090"
21..1454
/gene="GLUTX1"
21..1454
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CA875719.1"
/db_xref="GI:7018607"
/translation="MSPEDPEPTQPLRPPEARPRGRVFLASFAALGPIINFGALGSSPAIPSPRLRTAPALRLGDNAASWFGAVVTLGAAGCILGMLIDRSGRLSLILCTVPEVFAVITTAARDVMMILGRLTLGACVAVSVAYVISEIAYPAVRGLISGCVOLMVTGILLAYAAQVLEWRMLAVIGCVPTLLMLKTYMPTPEFLITQHOGEA MAALRFLMSGEQMEEPVGAHEQGFOLALLRRGYTKPLIGISLWFOQLSGVNAI MFYANSTFEFAKFKDSSLASVAVGIIQVLETAVALIMDRAGRLLALLAGVAVMFSM SAFCTYFKLQSLPSNNSHVLPIAEPYDVGVGLAMLVGSKCLEIAGFAVGMGPI PMLMSEIFPLHVKGVATGICVLINMFMAALVTRKEFNSVEMLRPYGAFMLTAAFCAL SVLETLIVEPETHGRILEQYTAHPEGR"
BASE COUNT 351 a 673 c 584 g 464 t
ORIGIN
Query Match 17.7%; Score 50; DB 10; Length 2072;
Best Local Similarity 100.0%; Pred. No. 8.5e-17;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 79 ggcctggagaccatccctggctcctcatgccaagatcctccctcagca 128
|||||
Db 1182 GCGTGGGAGCCATCCCTGCTCTCATGTCAGAGATCTTCCCTGCA 1231
RESULT 8
AF321324 1012 bp mRNA MAM 02-JUL-2001
LOCUS Bos taurus glucose transporter 8 (Glut8) mRNA, partial cds.
DEFINITION AF321324
ACCESSION AF321324
VERSION AF321324.1 GI:14582715
KEYWORDS
SOURCE
ORGANISM Bos taurus
COW.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 1012)
AUTHORS Augustin, R., Navarrete-Santos, A. and Fischer, B.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2000) Anatomy and Cell Biology, Faculty of Medicine, Martin Luther University Halle, Grosse Steinstrasse 52, Halle 061097, Germany
FEATURES Location/Qualifiers

source 1..1012
/organism="Bos taurus"
/db_xref="taxon:9913"
<1..1012
/gene="Glut8"
<1..1005
CDS
/gene="Glut8"
/codon_start=1
/product="glucose transporter 8"
/protein_id="AAK69606.1"
/db_xref="GI:14582716"
/translation="ISEIAYPEVRGLISCVOLMVTGILLAYLAGVLEWRMLAVLG CVAPSPMLLMCPMPTPEFLPSQHKHOEAMAMQFLMGYAQGEPEPPLGAQHDPHV AOLRRGVYKPIIGISLMAFQOLSGVNAVWFYAEYIFPEAKFKDSSLASVAVGIQV LPTATVALIMDRAGRLLTLTSGVAVFSTASGTVEFKLEGSPSSSHVDLPALVSM EADTNYGLAMLVAGNMKCLEIAGFAVGMGPIPLMLMSEIFPLHVKGVATGVTMPF MAFLVTRKEFSSLMELVRPYGAFMLASAFCTFYGLPTLACVPETKXGTLTDITAHPEGR"
BASE COUNT 153 a 323 c 302 g 234 t
ORIGIN
Query Match 15.6%; Score 44; DB 4; Length 1012;
Best Local Similarity 100.0%; Pred. No. 2e-13;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 157 gtcctcaacactggtcatgctctcgtgagcaagaagtt 200
|||||
Db 811 GTCTCACCACTGCTTCATGCGCTTCTGTGTCACCAAGATT 854
RESULT 9
HSA17801 1508 bp mRNA PRI 13-MAY-2000
LOCUS Homo sapiens mRNA for glucose transporter 8 (GLUT8 gene).
DEFINITION Y17801
ACCESSION Y17801
VERSION Y17801.1 GI:7688145
KEYWORDS glucose transporter 8; GLUT8 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1508)
AUTHORS Doege, H., Schurmann, A., Bahrenberg, G., Brauers, A. and Joost, H.G.
TITLE GLUT8, a novel member of the sugar transport facilitator family with glucose transport activity
JOURNAL J. Biol. Chem. 275 (21), 16275-16280 (2000)
MEDLINE 20283667
REFERENCE 2 (bases 1 to 1508)
AUTHORS Joost, H.G.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1998) H. Joost, Institute of Pharmacology and Toxicology, Technical University Aachen, Wendlingweg 2, D-52057 Aachen, FRG
FEATURES Location/Qualifiers
source 1..1508
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="testis"
27..1460
/gene="GLUT8"
27..1460
CDS
/gene="GLUT8"
/codon_start=1
/product="glucose transporter 8"
/protein_id="CA889809.1"
/db_xref="GI:7688146"
/translation="MTPEDPEPTQPLRPGSGAPRGVFLAFAAALGPIISFGFAL GYSSPAIPSLQRAAPAPRLDDAASWFGAVVTLGAAGCVLGMIVDRAGRKLSILL GYSPFAVFAVITTAADVMMILGRLTLGACVAVSVAYVISEIAYPAVRGLISGC VOLMVTGILLAYLAGVLEWRMLAVIGCVPTLLMLKTYMPTPEFLITQHOGEA MAALRFLMSGEQMEEPVGAHEQGFOLALLRRGYTKPLIGISLWFOQLSGVNAI MFYANSTFEFAKFKDSSLASVAVGIIQVLETAVALIMDRAGRLLALLAGVAVMFSM SAFCTYFKLQSLPSNNSHVLPIAEPYDVGVGLAMLVGSKCLEIAGFAVGMGPI PMLMSEIFPLHVKGVATGICVLINMFMAALVTRKEFNSVEMLRPYGAFMLTAAFCAL SVLETLIVEPETHGRILEQYTAHPEGR"
BASE COUNT 153 a 323 c 302 g 234 t
ORIGIN

AFGAYFKLTGGPGNSHVAISAPVADPDAVGLAMVAGMCLFTAGFAVGMGPI
 PMLMSEIFPLHVKGVATGICVLTFTNMLAFVTKFSSLMELVRYPGAFMLASAFICF
 SVLTFECVPEIKKTLQITAHFEGR"
 BASE COUNT 194 a 519 c 477 g 318 t
 ORIGIN

Query Match 14.5%; Score 41; DB 9; Length 1508;
 Best Local Similarity 100.0%; Pred. No. 9,3e-12;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 88 cccatccctgctcctcatgctcagagatcttcctctgca 128
 ||||||||||||||||||||||||||||||||||||||||
 Db 1197 CCCATCCCTGGCTCCTCATGTCAAGAGATCTTCCCTGCA 1237

RESULT 10
 HSA245937 1873 bp mRNA PRI 18-FEB-2000
 LOCUS Homo sapiens mRNA for glucose transporter (GLUTX1 gene).
 DEFINITION AJ245937
 VERSION AJ245937.1 GI:7018305
 KEYWORDS glucose transporter; GLUTX1 gene.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1873)
 AUTHORS Ibberson,M., Uldry,M. and Thorens,B.
 TITLE GLUTX1, a novel mammalian glucose transporter expressed in the central nervous system and insulin-sensitive tissues
 JOURNAL J. Biol. Chem. 275 (7), 4607-4612 (2000)
 MEDLINE 20138191
 REFERENCE 2 (bases 1 to 1873)
 AUTHORS Ibberson,M.R.
 TITLE Direct Submission
 JOURNAL Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne, CH-1005, SWITZERLAND

FEATURES
 Source Location/Qualifiers

1..1873
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 4..1437
 /gene="GLUTX1"
 4..1437
 /gene="GLUTX1"
 /codon_start=1
 /product="glucose transporter"
 /protein_id="CAB75702.1"
 /db_xref="GI:7018305"

/translation="MPEDPEETQPLGPPGSAAPGRRVFLAFAAALGPLSGFAL
 GYSSPAIPSLORAAPAPRLDDAASMGAVVTIGAAAGVGLGMLVDRAGRKLSLL
 CSVPFVAGFAVITAQDVWMLGRLTLGLACGASVLAIVYISEIAYPAVRGLLGS
 VOLWVVGILAYLAGVLEWRLAVLCVPSLMLLMCPETPRLLTQHRQEA
 MAALRFLMGSEQGWEDPPIGAEOSEFLALRQPGIYKPFILIGSLAQOQSGVNAV
 FYAETIIEEAKFKDSSLASVYGVIOVFTFAVALIMDRAGRLLVLGCVVVFST
 AFGAYFKLTGGPGNSHVAISAPVADPDAVGLAMVAGMCLFTAGFAVGMGPI
 PMLMSEIFPLHVKGVATGICVLTFTNMLAFVTKFSSLMELVRYPGAFMLASAFICF
 SVLTFECVPEIKKTLQITAHFEGR"
 BASE COUNT 279 a 621 c 573 g 400 t
 ORIGIN

Query Match 14.5%; Score 41; DB 9; Length 1873;
 Best Local Similarity 100.0%; Pred. No. 9,2e-12;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 88 cccatccctgctcctcatgctcagagatcttcctctgca 128
 ||||||||||||||||||||||||||||||||||||||||
 Db 1174 CCCATCCCTGGCTCCTCATGTCAAGATCTTCCCTGCA 1214

RESULT 11
 AX179740 2080 bp DNA PAT 06-AUG-2001
 LOCUS Sequence 28 from Patent WO0146258.
 DEFINITION AX179740
 ACCESSION AX179740
 VERSION AX179740.1 GI:15132104
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2080)
 AUTHORS Baughn,M.R., Burford,N., Au-Young,J., Lu,D.A., Yang,J., Reddy,R.,
 Lal,P., Hillman,J.L., Azimzai,Y., Yue,H., Nguyen,D.B., Yao,M.G.,
 Gandhi,A.R., Tang,Y.T. and Khan,F.A.
 TITLE Transporters and ion channels
 JOURNAL Patent: WO 0146258-A 28-28-JUN-2001;
 Incyte Genomics, Inc. (US)

FEATURES
 source Location/Qualifiers
 1..2080
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 1416107CB1"
 BASE COUNT 300 a 688 c 632 g 460 t
 ORIGIN

Query Match 14.5%; Score 41; DB 6; Length 2080;
 Best Local Similarity 100.0%; Pred. No. 9,1e-12;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 88 cccatccctgctcctcatgctcagagatcttcctctgca 128
 ||||||||||||||||||||||||||||||||||||||||
 Db 1216 CCCATCCCTGGCTCCTCATGTCAAGATCTTCCCTGCA 1256

RESULT 12
 AX076667 2217 bp DNA PAT 06-FEB-2001
 LOCUS Sequence 1 from Patent WO0104145.
 DEFINITION AX076667
 ACCESSION AX076667
 VERSION AX076667.1 GI:12711198
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2217)
 AUTHORS Thorens,B., Ibberson,M. and Uldry,M.
 TITLE Glutx polypeptide family and nucleic acids encoding same
 JOURNAL Patent: WO 0104145-A 1 18-JAN-2001;
 University of Lausanne (CH)

FEATURES
 source Location/Qualifiers
 1..2217
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 348..1781
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC28495.1"
 /db_xref="GI:12711198"

CDS

/translation="MPEDPEETQPLGPPGSAAPGRRVFLAFAAALGPLSGFAL
 GYSSPAIPSLORAAPAPRLDDAASMGAVVTIGAAAGVGLGMLVDRAGRKLSLL
 CSVPFVAGFAVITAQDVWMLGRLTLGLACGASVLAIVYISEIAYPAVRGLLGS
 VOLWVVGILAYLAGVLEWRLAVLCVPSLMLLMCPETPRLLTQHRQEA
 MAALRFLMGSEQGWEDPPIGAEOSEFLALRQPGIYKPFILIGSLAQOQSGVNAV
 FYAETIIEEAKFKDSSLASVYGVIOVFTFAVALIMDRAGRLLVLGCVVVFST
 AFGAYFKLTGGPGNSHVAISAPVADPDAVGLAMVAGMCLFTAGFAVGMGPI
 PMLMSEIFPLHVKGVATGICVLTFTNMLAFVTKFSSLMELVRYPGAFMLASAFICF
 SVLTFECVPEIKKTLQITAHFEGR"
 BASE COUNT 320 a 750 c 672 g 475 t
 ORIGIN

Query Match 14.5%: Score 41; DB 6; Length 2217;
Best Local Similarity 100.0%; Pred. No. 9,1e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 88 cccatccctgctccatgctcagagatcttccctgcga 128
|||||
Db 1518 CCCATCCCTGCTCCATGTCAGAGATCTTCCCTCGCA 1558

RESULT 13

LOCUS AL445222 225370 bp DNA PRI 24-APR-2001
DEFINITION Human DNA sequence from clone RP11-356b19 on chromosome 9, complete sequence.
ACCESSION AL445222
VERSION AL445222
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 225370)
AUTHORS Laird, G.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2001) Sanger Centre, Hinxton, Cambridgehire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
Requests: clonerequests@sanger.ac.uk
On Apr 26, 2001 this sequence version replaced gi:13277497.

COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; SW., SWISSPROT; Tr., TrEMBL; MP., MOPREP; Information on the MOPREP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-356b19 is from the library RPCT-11.2 constructed by the group of Plieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
This sequence is the entire insert of clone RP11-356b19. The true left end of clone RP11-373J8 is at 96439 in this sequence. The true right end of clone RP11-225021 is at 9980 in this sequence.

FEATURES

SOURCE
1..225370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-356b19"
/clone_lib="RPCT-11.2"
23455..23821
/note="Single clone region. Assembly confirmed by restriction digest data."
65019..65283
/note="Single clone region. Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."
225123..225370
misc_feature
misc_feature
misc_feature

/note="Single clone region. Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."
BASE COUNT 56329 a 57092 c 57267 g 54682 t
ORIGIN

Query Match 14.5%: Score 41; DB 9; Length 225370;
Best Local Similarity 100.0%; Pred. No. 7e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 88 cccatccctgctccatgctcagagatcttccctgcga 128
|||||
Db 61234 CCCATCCCTGCTCCATGTCAGAGATCTTCCCTCGCA 61274

RESULT 14

LOCUS AX191507 1461 bp DNA PAT 15-AUG-2001
DEFINITION Sequence 29 from Patent WO0149728.
ACCESSION AX191507
VERSION AX191507.1 GI:15209697
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1461)
AUTHORS Kato, S. and Kimura, T.
TITLE Human proteins having hydrophobic domains and dnas encoding these proteins
JOURNAL Patent: WO 0149728-A 29 12-JUL-2001;
Protegene Inc. (JP) : SAGAMI CHEMICAL RESEARCH CENTER (JP)
FEATURES
source
location/Qualifiers
1..1461
/organism="Homo sapiens"
/db_xref="taxon:9606"
61..849
/note="unnamed protein product"
/codon_start=1
/protein_id="CA51153.1"
/db_xref="GI:15209698"
/translation="MPPEDPEEPQPLGPPGSAAPKGRVFLAFAAAGLPSFGFAL
GYSPALPSLQRAAPAPRLDDAASWFGAVVVLGAAGVIGWLVDRGRSLSL
CYPFVAGFAVITAAQDVMLLGLRLTGACVASIVAFVYSEIAYPAVRGLLSC
VOLMVVGLLAVLAGVLEWRVLAVGCVPSIMLLMFPEETRPFLTLQRRQEA
APGIVRGHGVQHECLRRLLQAPRQPMQALLARGHGLGACLTAC"

CDS

BASE COUNT 195 a 501 c 456 g 309 t
ORIGIN

Query Match 8.2%: Score 23; DB 6; Length 1461;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 ggcagcatgctcccttcacgc 65
|||||
Db 881 GGCAGCATGCTCCCTTCATGCC 903

RESULT 15

LOCUS CGRHOD 11931 bp DNA ROD 27-MAY-1992
DEFINITION C. griseus rhodopsin gene for opsin protein.
ACCESSION X61084
VERSION X61084.1 GI:49478
KEYWORDS opsin; rhodopsin; vision protein.
SOURCE Chinese hamster.
ORGANISM Cricetus griseus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetus.

REFERENCE 1 (bases 1 to 11931)

AUTHORS Gale, J.M., Tobey, R.A. and D'Anna, J.A.
 TITLE Localization and DNA sequence of a replication origin in the
 rhodopsin gene locus of Chinese hamster cells
 JOURNAL J. Mol. Biol. 224 (2), 343-358 (1992)
 MEDLINE 92219256
 REFERENCE 2 (bases 1 to 11931)
 AUTHORS D'Anna, A.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUL-1991) A. D'Anna, Los Alamos National Lab, PO Box
 1663/MS-M880, Los Alamos, NM 87545, USA

FEATURES
 source
 1..11931
 /organism="Crictetus griseus"
 /db_xref="taxon:10029"
 /tissue_type="ovary"
 /cell_line="CHO"
 /clone_lib="cosmid: SCOS1"
 3269..3426
 join(6107..6467,7929..8097,9177..9342,9455..9694,
 10702..10812)
 /gene="rhodopsin"
 6107..10812
 /gene="rhodopsin"
 join(6107..6467,7929..8097,9177..9342,9455..9694,
 10702..10812)
 /gene="rhodopsin"
 /codon_start=1
 /product="opsin"
 /protein_id="CAA43398.1"
 /db_xref="GI:48479"
 /translation="MNGTEGPNFYVPEFNSNAGVRSPEYPOYLAEPMQSMLAAYM
 FLILVIGFPINELTLYTVQHKLRTPLYILNLAVADLFVFGFTTLYTSLHGY
 FVEGPTGCLKEGFATLGLTALMSLVALTERVYVICKPMSNFERGENHAIMGVET
 MIMACAPPLVIGMSRYIPGMOCSGVDYTLKPEVNNESFVYIMVFHTIPLIV
 IFFCGQLVFTYKEAAQQQESATTQKAEKVTNVIIMVFFELICMPYACVAFYIF
 THQSNFGPIFTLDPAFKSSSYINPVYIMMNRQFRNCMLTLTLCGKNILGDDEAS
 AVASKTETSQVAPA"
 prln_transcript
 <6107..>10812
 /gene="rhodopsin"
 join(<6107..6467,7929..8097,9177..9342,9455..9694,
 10702..>10812)
 /gene="rhodopsin"
 <6107..6467
 /gene="rhodopsin"
 /number=1
 6468..7928
 /gene="rhodopsin"
 /number=1
 6944..7070
 /rpt_family="Alu"
 7929..8097
 /gene="rhodopsin"
 /number=2
 8098..9176
 /gene="rhodopsin"
 /number=2
 9177..9342
 /gene="rhodopsin"
 /number=3
 9343..9454
 /gene="rhodopsin"
 /number=3
 9455..9694
 /gene="rhodopsin"
 /number=4
 9695..10701
 /gene="rhodopsin"
 /number=4
 10031..10159
 /rpt_family="Alu"
 10702..>10812
 /gene="rhodopsin"

repeat_region 11409..11462 /number=5
 /note="triple motif (GA)"
 BASE COUNT 2719 a 3172 c 3122 g 2918 t
 ORIGIN

Query Match 7.4%; Score 21; DB 10; Length 11931;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 32 ggcctgagctgtaggcagcatgt 52
 |||||
 Db 591 GCCTGCTGTAGCAGCATGT 611

Search completed: February 13, 2002, 20:07:08
 Job time: 23045 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 20:10:32 ; Search time 282.42 Seconds
(without alignments)
856.050 Million cell updates/sec

Title: US-09-516-493-11

Perfect score: 282

Sequence: 1 gagcctgctgattcaccct.....tcctcagcgtccttcacg 282

Scoring table: OLIGO.NUC
Gapop 60.0 , Gapext 60.0

Searched: 930621 seqs, 428662619 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

N_Geneseq_1101.*
1: /SID52/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	282	100.0	2087	22	AAF55866 Rat GLUTX1 coding
2	50	17.7	2072	22	AAF55867 Murine GLUTX1 codi
3	41	14.5	579	21	AAA44914 Human secreted exp
4	41	14.5	2080	22	AAD09552 Human transporter
5	41	14.5	2217	22	AAF53865 Human GLUTX1 codin
6	23	8.2	1461	22	AAID2574 Human protein havi
7	20	7.1	423	22	AAI14336 Probe #4269 for ge
8	20	7.1	423	22	AAI35713 Probe #4399 used t
9	20	7.1	423	22	AAI04170 Probe #4161 used t
10	20	7.1	465	22	AAI23539 Probe #13472 for g
11	20	7.1	465	22	AAI48856 Probe #17342 used

C	12	20	7.1	465	22	AAI09165	Probe #9156 used t
C	13	19	6.7	954	18	AAAX30839	Streptococcus pneu
C	14	19	6.7	2051	19	AAV373386	Streptococcus pneu
C	15	19	6.7	4214	22	AAH573735	Human heart cell s
C	16	19	6.7	10399	19	AAV522933	Streptococcus pneu
C	17	18	6.4	1058	19	AAI986661	DNA encoding a S.
C	18	18	6.4	2831	19	AAV52942	Rat UNC-5 homology
C	19	18	6.4	10442	22	AAV24680	Nucleotide sequenc
C	20	18	6.4	10442	22	AAE24702	Nucleotide sequenc
C	21	18	6.4	10474	22	AAE24685	Nucleotide sequenc
C	22	18	6.4	10474	22	AAE24685	Nucleotide sequenc
C	23	18	6.4	10474	22	AAE24685	Nucleotide sequenc
C	24	18	6.4	10474	22	AAE24685	Nucleotide sequenc
C	25	17	6.0	160	22	AAI27584	Probe #17517 for g
C	26	17	6.0	160	22	AAI56555	Probe #25241 used
C	27	17	6.0	393	22	AAE64435	Novel human polynu
C	28	17	6.0	399	22	AAE07633	Human secreted pro
C	29	17	6.0	421	22	AAE64504	Novel human polynu
C	30	17	6.0	573	22	AAI18529	Probe #8462 for ge
C	31	17	6.0	573	22	AAI143648	Probe #12334 used
C	32	17	6.0	628	21	AAE07821	Fusarium venenatum
C	33	17	6.0	677	21	AAE45084	Arabidopsis thalia
C	34	17	6.0	868	22	AAH03721	Human CDNA clone (
C	35	17	6.0	879	22	AAH53710	S. epidermidis ope
C	36	17	6.0	879	22	AAH53820	S. epidermidis ope
C	37	17	6.0	1497	21	AAA15943	Human protein clon
C	38	17	6.0	1498	21	AAA37771	Human TWIK-2 codin
C	39	17	6.0	1767	22	AAH13832	Human CDNA sequenc
C	40	17	6.0	1971	22	AAH160781	Human polynucleoti
C	41	17	6.0	1986	22	AAI58995	Human polynucleoti
C	42	17	6.0	2848	22	AAH16234	Human CDNA sequenc
C	43	17	6.0	2913	22	AAH54542	S. epidermidis gen
C	44	17	6.0	3150	22	AAH54334	S. epidermidis gen
C	45	17	6.0	3452	21	AAA37770	Human TWIK-2 gene.
C	46	17	6.0	3452	22	AAE04872	Human TWIK-2 CDNA
C	47	17	6.0	3514	21	AAH27747	Human potassiu ch
C	48	17	6.0	3589	22	AAH53989	S. epidermidis gen
C	49	17	6.0	3768	21	AAH15953	Human protein clon
C	50	17	6.0	4198	22	AAH54240	S. epidermidis gen
C	51	17	6.0	4360	22	AAE09494	Human SGP18 phosph
C	52	16	5.7	21	18	AAE45264	Heregulin reverse
C	53	16	5.7	241	14	AAE60561	Human brain Expres
C	54	16	5.7	318	21	AAE00716	Human secreted pro
C	55	16	5.7	329	21	AAE17356	Human secreted pro
C	56	16	5.7	376	14	AAE09649	Human brain Expres
C	57	16	5.7	390	22	AAE67624	C glutamic codin
C	58	16	5.7	438	22	AAH94142	Human foetal CDNA.
C	59	16	5.7	475	21	AAE79720	Human secreted pro
C	60	16	5.7	476	21	AAE00716	Human secreted pro
C	61	16	5.7	512	21	AAE00713	Human secreted pro
C	62	16	5.7	550	20	AAE207225	Human lung tumour
C	63	16	5.7	550	21	AAE79115	Human lung tumour
C	64	16	5.7	559	15	AAE66938	GGF gene coding se
C	65	16	5.7	569	15	AAE67003	GGF gene coding se
C	66	16	5.7	569	16	AAE74906	Human glial cell
C	67	16	5.7	569	17	AAE13855	Human glial growth
C	68	16	5.7	569	17	AAE30984	Human glial growth
C	69	16	5.7	569	17	AAE06725	Human glial growth
C	70	16	5.7	569	16	AAE81194	Human neuregulin g
C	71	16	5.7	580	22	AAI18128	Probe #8061 for ge
C	72	16	5.7	580	22	AAI43134	Probe #11820 used
C	73	16	5.7	597	21	AAE76053	Human ORFX CR930.
C	74	16	5.7	621	20	AAE89698	Human collagen gen
C	75	16	5.7	631	22	AAE21955	Human collagen gen
C	76	16	5.7	657	21	AAE79872	Human secreted pro
C	77	16	5.7	777	22	AAH03917	Human CDNA clone (
C	78	16	5.7	778	16	AAE80221	Human PROMDF-beta2
C	79	16	5.7	897	22	AAE80050	Nucleotide sequenc
C	80	16	5.7	959	14	AAE49499	Nucleotide sequenc
C	81	16	5.7	959	17	AAI12944	Tsxex. Meleagris g
C	82	16	5.7	959	19	AAE06455	Short Tsex sequenc
C	83	16	5.7	1063	16	AAE80215	Human NDF-alpha
C	84	16	5.7	1110	22	AAH34059	Human colon cancer

85	16	5.7	1184	20	AAZ17768	Human gene express	c 158	16	5.7	5270	13	AAQ25387	Rat thymotropin re
c 86	16	5.7	1261	16	AAQ80230	Rat NDF clone 40 D	c 159	16	5.7	6202	22	AAI2940	Human genomic DNA
c 87	16	5.7	1328	21	AAC59468	Human secreted pro	c 160	16	5.7	6880	22	AAI294734	Human ATP binding
c 88	16	5.7	1351	22	AAC84374	Nucleotide sequenc	c 161	16	5.7	6928	11	AAO06161	Part 1 of cDNA des
c 89	16	5.7	1369	22	AAI57925	Human polynucleoti	c 162	16	5.7	7857	27	AAI57925	Human ABC1 cholest
c 90	16	5.7	1403	22	AAC60224	Human hydrolase-1i	c 163	16	5.7	7860	22	AAI57925	Human ABC1 nucleot
c 91	16	5.7	1404	22	AAI59711	Human polynucleoti	c 164	16	5.7	7860	22	AAI57925	Human ABC1 cholest
c 92	16	5.7	1446	22	AAC84375	CPD domain of huma	c 165	16	5.7	7861	21	AAI57925	Human ABC1 cholest
c 93	16	5.7	1453	22	AAH13841	Human cDNA sequenc	c 166	16	5.7	7864	21	AAI57925	Human ABC1 cholest
c 94	16	5.7	1637	20	AAH90600	Mouse aminopeptida	c 167	16	5.7	7864	21	AAI57925	Human ABC1 cholest
c 95	16	5.7	1651	16	AAO80218	Human NDF-alpha2b	c 168	16	5.7	7864	21	AAI57925	Human ABC1 cholest
c 96	16	5.7	1714	22	AAH98311	Human EST-derived	c 169	16	5.7	7864	21	AAI57925	Human ABC1 cholest
c 97	16	5.7	1752	22	AAI57408	Human p95 protein	c 170	16	5.7	9181	22	AAI57925	Human ABC1 cholest
c 98	16	5.7	1807	16	AAO80217	Human NDF-alpha2b	c 171	16	5.7	9741	22	AAI57925	Human ABC1 cholest
c 99	16	5.7	1873	16	AAO80220	Human NDF-beta1a c	c 172	16	5.7	9854	22	AAI57925	Human ABC1 cholest
c 100	16	5.7	1894	14	AAO61796	Human NRSE cDNA.	c 173	16	5.7	20084	22	AAI57925	Human ABC1 cholest
c 101	16	5.7	1894	16	AAO80224	Rat NDF clone 44 D	c 174	16	5.7	20084	22	AAI57925	Human ABC1 cholest
c 102	16	5.7	1895	21	AAI14638	Aspergillus oryzae	c 175	16	5.7	20284	21	AAI57925	Human ABC1 cholest
c 103	16	5.7	1897	21	AAC98161	Human colon cancer	c 176	16	5.7	24053	22	AAI12308	Human ABC1 cholest
c 104	16	5.7	1899	19	AAV06456	Extended rtxex sequ	c 177	16	5.7	26698	17	AAI08126	Mouse syndecan-1 g
c 105	16	5.7	1991	22	AAH50967	Alpha carboxyltran	c 178	16	5.7	26700	15	AAO67902	Syndecan gene. Mu
c 106	16	5.7	2010	13	AAO31541	Human heregulin-al	c 179	16	5.7	26700	19	AAV15946	Mouse syndecan gen
c 107	16	5.7	2010	22	AAD00786	Human heregulin-HR	c 180	16	5.7	26700	20	AAV15946	Mouse syndecan-1 e
c 108	16	5.7	2011	22	AAI55871	Rat GLUTR3 coding	c 181	16	5.7	32127	22	AAI53605	Human kidney relat
c 109	16	5.7	2029	13	AAO31542	Human heregulin be	c 182	16	5.7	32207	20	AAI53605	KSHV LTR DNA (nucl
c 110	16	5.7	2164	18	AAI58558	Human breast cance	c 183	16	5.7	65792	22	AAI28544	Genomic fragment #
c 111	16	5.7	2164	20	AAI58558	Full cDNA sequenc	c 184	16	5.7	65792	22	AAI28544	KSHV long unique c
c 112	16	5.7	2164	21	AAO72221	Full length gp30/a	c 185	16	5.7	183999	22	AAI28544	Human ABC1 genomic
c 113	16	5.7	2186	14	AAO51344	Rat NRSE cDNA.	c 186	16	5.7	349980	22	AAI28544	C glutamic codin
c 114	16	5.7	2186	16	AAO80223	Rat NDF DNA. Rat	c 187	16	5.3	27	19	AAI28544	Human G-protein co
c 115	16	5.7	2187	22	AAI5893	Human cDNA sequenc	c 188	16	5.3	27	19	AAI28544	Primer HT5r-2 for
c 116	16	5.7	2196	22	AAH15819	Human cDNA sequenc	c 189	16	5.3	27	19	AAI28544	Primer HT5r-2 for
c 117	16	5.7	2199	18	AAI58559	Human breast cance	c 190	16	5.3	27	19	AAI28544	PCR primer HT5r-2
c 118	16	5.7	2199	20	AAI587701	Human heregulin-b	c 191	16	5.3	32	18	AAI28544	Primer for CCR4 CC
c 119	16	5.7	2199	20	AAI587701	Full cDNA sequenc	c 192	16	5.3	38	21	AAI28544	Human CCR4 PCR pri
c 120	16	5.7	2199	21	AAD00787	Human heregulin-HR	c 193	16	5.3	85	22	AAI28544	Probe #12821 for g
c 121	16	5.7	2199	21	AAO72222	Full length gp30/b	c 194	16	5.3	85	22	AAI28544	Probe #16873 used
c 122	16	5.7	2226	20	AAI587700	Human heregulin-al	c 195	16	5.3	85	22	AAI28544	Probe #546 used t
c 123	16	5.7	2226	21	AAD00816	Human heregulin-HR	c 196	16	5.3	136	22	AAI28544	Probe #23511 used
c 124	16	5.7	2274	19	AAV58359	Coding sequence fo	c 197	16	5.3	144	16	AAI28544	Human gene signatu
c 125	16	5.7	2277	22	AAI14628	Human cDNA sequenc	c 198	16	5.3	148	16	AAI28544	Human cervical can
c 126	16	5.7	2288	22	AAI60582	Human polynucleoti	c 199	16	5.3	165	21	AAI17778	Aspergillus niger
c 127	16	5.7	2315	22	AAH15989	Human cDNA sequenc	c 200	16	5.3	168	22	AAI28544	Probe #23523 used
c 128	16	5.7	2335	16	AAO80216	Human pRONDF-alpha	c 201	16	5.3	188	22	AAI28544	Probe #17330 for g
c 129	16	5.7	2351	15	AAO62284	Neurotrophic facto	c 202	16	5.3	188	22	AAI28544	Probe #24992 used
c 130	16	5.7	2351	18	AAI588210	cDNA for chicken n	c 203	16	5.3	195	22	AAI28544	Probe #11847 for g
c 131	16	5.7	2356	16	AAO80229	Rat NDF clone 38 D	c 204	16	5.3	195	22	AAI28544	Probe #15885 used
c 132	16	5.7	2430	16	AAO80232	Rat NDF clone 42A	c 205	16	5.3	195	22	AAI28544	Probe #7593 used t
c 133	16	5.7	2431	13	AAO31546	Human heregulin-b	c 206	16	5.3	202	21	AAI28544	Human secreted exp
c 134	16	5.7	2431	20	AAO87704	Human heregulin-b	c 207	16	5.3	219	21	AAI28544	Human secreted pro
c 135	16	5.7	2431	21	AAD00790	Human heregulin-HR	c 208	16	5.3	227	16	AAI28544	Human gene signatu
c 136	16	5.7	2461	22	AAI60171	Human polynucleoti	c 209	16	5.3	286	21	AAI28544	Human secreted pro
c 137	16	5.7	2490	13	AAO31544	Human heregulin-b	c 210	16	5.3	287	21	AAI28544	Human colon cancer
c 138	16	5.7	2490	20	AAI587702	Human heregulin-b	c 211	16	5.3	342	17	AAI28544	G-protein conjugat
c 139	16	5.7	2490	21	AAD00788	Human heregulin-HR	c 212	16	5.3	347	17	AAI28544	Human lung tumour
c 140	16	5.7	2531	16	AAO80227	Rat NDF clone 20 D	c 213	16	5.3	367	16	AAI28544	Human gene signatu
c 141	16	5.7	2540	22	AAI580049	Nucleotide sequenc	c 214	16	5.3	378	22	AAI28544	Novel human polynu
c 142	16	5.7	2545	11	AAO04645	Encodes beta subun	c 215	16	5.3	384	22	AAI28544	Nucleotide sequenc
c 143	16	5.7	2545	14	AAO14734	Encodes beta subun	c 216	16	5.3	390	21	AAI28544	Human secreted pro
c 144	16	5.7	2545	14	AAO51021	Human FeoRT beta g	c 217	16	5.3	390	22	AAI28544	Human secreted pro
c 145	16	5.7	2580	18	AAI586009	Mouse gamma melfri	c 218	16	5.3	398	22	AAI28544	Probe #104 for gen
c 146	16	5.7	2805	22	AAH17148	Rat NDF clone 41 D	c 219	16	5.3	398	22	AAI28544	Probe #107 used to
c 147	16	5.7	2805	22	AAH17148	Human cDNA sequenc	c 220	16	5.3	398	22	AAI28544	Probe #105 used to
c 148	16	5.7	2856	22	AAI580048	Nucleotide sequenc	c 221	16	5.3	406	21	AAI28544	Human secreted pro
c 149	16	5.7	2862	22	AAH14361	Human cDNA sequenc	c 222	16	5.3	414	21	AAI28544	Human colon cancer
c 150	16	5.7	2914	16	AAO80226	Rat NDF clone 19 D	c 223	16	5.3	424	21	AAI28544	Human secreted pro
c 151	16	5.7	3086	22	AAI580059	Nucleotide sequenc	c 224	16	5.3	437	22	AAI28544	Human secreted pro
c 152	16	5.7	3161	16	AAO80233	Rat NDF clone 42B	c 225	16	5.3	437	22	AAI28544	Probe #3600 for ge
c 153	16	5.7	3212	22	AAO84373	Nucleotide sequenc	c 226	16	5.3	437	22	AAI28544	Probe #3712 used t
c 154	16	5.7	3344	16	AAO80228	Rat NDF clone 22 D	c 227	16	5.3	437	22	AAI28544	Probe #3540 used t
c 155	16	5.7	4332	19	AAV31717	Kaposi's sarcoma a	c 228	16	5.3	456	22	AAI28544	Nucleotide sequenc
c 156	16	5.7	5143	20	AAI20131	Human CDK-inhibito	c 229	16	5.3	458	22	AAI28544	Probe #17 for gene
c 157	16	5.7	5143	20	AAI15110	Transcriptional re	c 230	16	5.3	458	22	AAI28544	Probe #18 used to

231	15	5.3	458	22	AA100025	Probe #16 used to	304	15	5.3	1050	22	AAE94441	Human hydrophobic
232	15	5.3	466	21	AAD00815	Human irritable bo	305	15	5.3	1053	22	AAE98857	Nucleotide sequenc
233	15	5.3	466	21	AA111221	Probe #1154 for ge	306	15	5.3	1058	21	AAE18331	Lung cancer associ
234	15	5.3	466	22	AA132483	Probe #1169 used t	307	15	5.3	1104	21	AAD01958	Murine TH2AF1 isot
235	15	5.3	466	22	AA101137	Probe #1128 used t	308	15	5.3	1108	20	AAE77557	Human ovarian tumo
236	15	5.3	470	22	AA112733	Probe #2666 for ge	309	15	5.3	1110	22	AA158834	Human polynucleoti
237	15	5.3	470	22	AA134084	Probe #2770 used t	310	15	5.3	1120	22	AA160620	Human polynucleoti
238	15	5.3	470	22	AA102642	Probe #2633 used t	311	15	5.3	1124	19	AAV07373	Mouse protease-act
239	15	5.3	471	22	AAE98856	Nucleotide sequenc	312	15	5.3	1126	21	AAC39072	Arabidopsis thalia
240	15	5.3	471	21	AAC00181	Human secreted pro	313	15	5.3	1136	19	AAV12306	Ribonucleotide red
241	15	5.3	484	22	AA112181	Probe #2114 for ge	314	15	5.3	1159	22	AAC91601	Human angiotensino
242	15	5.3	484	22	AA103527	Probe #2213 used t	315	15	5.3	1175	20	AAAS2218	Protein PRO187 CDN
243	15	5.3	484	22	AA102094	Probe #2085 used t	316	15	5.3	1175	20	AAAX8429	FGF-8 homologue PR
244	15	5.3	486	22	AAH62266	Shrimp white spot	317	15	5.3	1185	19	AAV43608	Human secreted pro
245	15	5.3	497	19	AAV33224	Fibroblast growth	318	15	5.3	1200	21	AAV77525	Human PRO187 CDNA
246	15	5.3	498	22	AAC90688	Human PCF-13 const	319	15	5.3	1200	21	AAA30024	Human PRO187 nucle
247	15	5.3	504	20	AAE20915	Nucleotide sequenc	320	15	5.3	1200	22	AAE21380	Human CDNA sequenc
248	15	5.3	507	22	AAE98855	Nucleotide sequenc	321	15	5.3	1200	22	AAE58494	PRO187 coding sequ
249	15	5.3	509	22	AA141790	Probe #10476 used t	322	15	5.3	1200	22	AAC87033	Nucleotide sequenc
250	15	5.3	512	22	AA140216	Probe #8902 used t	323	15	5.3	1200	22	AAE72376	Human PRO187 CDNA
251	15	5.3	514	17	AAE35278	Chemokine receptor	324	15	5.3	1210	18	AAE95246	Human IA-2/ICA512
252	15	5.3	524	19	AAV33225	Fibroblast growth	325	15	5.3	1212	22	AAC90672	Human fibroblast g
253	15	5.3	524	22	AA118265	Probe #8198 for ge	326	15	5.3	1230	8	AAE70677	Sequence encoding
254	15	5.3	524	22	AA143321	Probe #12007 used	327	15	5.3	1230	8	AAE70053	Benzene oxygenase
255	15	5.3	525	8	AAE70146	Human FGF-13 const	328	15	5.3	1230	8	AAE70822	DNA encoding enzym
256	15	5.3	525	22	AAC90691	Human colon cancer	329	15	5.3	1235	19	AAE11726	Human G-protein co
257	15	5.3	531	22	AAH36467	Human cervical can	330	15	5.3	1235	19	AAE89983	Human G-protein co
258	15	5.3	539	22	AAH68914	Human G-protein can	331	15	5.3	1235	19	AAE94184	Human G-protein co
259	15	5.3	541	22	AAH06067	Human CDNA clone (332	15	5.3	1235	20	AAE89527	CDNA for a human G
260	15	5.3	541	22	AAH17299	Probe #7232 for ge	333	15	5.3	1254	16	AAE00621	Human (2'-5') olig
261	15	5.3	543	22	AA142188	Probe #10874 used	334	15	5.3	1254	22	AAH68181	Sequence of (2'-5'
262	15	5.3	543	22	AAH72584	Human cervical can	335	15	5.3	1299	22	AAH68181	Sequence of (2'-5'
263	15	5.3	545	22	AAH03206	Human CDNA clone (336	15	5.3	1322	16	AAE61024	Human (2'-5') olig
264	15	5.3	547	22	AAH09569	Human CDNA clone (337	15	5.3	1341	20	AAE61024	P. putida KT2440-a
265	15	5.3	554	19	AAV33226	Fibroblast growth	338	15	5.3	1347	20	AAE33943	Human HCMV inducib
266	15	5.3	558	22	AAC90694	Human PCF-13 const	339	15	5.3	1347	21	AAE338325	Human angiotensino
267	15	5.3	561	22	AAE93563	CDNA encoding SRT	340	15	5.3	1350	20	AAE200364	Nucleotide sequenc
268	15	5.3	572	22	AAH12219	Human CDNA clone (341	15	5.3	1350	22	AAE89026	Rat PAP4 partial
269	15	5.3	576	22	AAH11802	Probe #10488 used	342	15	5.3	1376	21	AAC45796	Arabidopsis thalia
270	15	5.3	584	22	AAH10043	Human CDNA clone (343	15	5.3	1381	20	AAE33469	Human prostate can
271	15	5.3	646	21	AAZ80574	Human colon cancer	344	15	5.3	1389	22	AAE88330	S. splosa DNA fra
272	15	5.3	671	17	AAE40818	Serine protease nt	345	15	5.3	1394	21	AAC77902	Human cancer assoc
273	15	5.3	672	22	AAC90823	Flca serine protea	346	15	5.3	1394	22	AAH33200	Human colon cancer
274	15	5.3	675	22	AAE97894	Human secreted pro	347	15	5.3	1405	22	AAH161232	Human polynucleoti
275	15	5.3	678	22	AA120443	Probe #10376 for g	348	15	5.3	1413	20	AAZ32341	Human pancreatic i
276	15	5.3	678	22	AA145646	Probe #14332 used	349	15	5.3	1413	20	AAZ09922	Human islet cell a
277	15	5.3	678	22	AA106142	Probe #6133 used t	350	15	5.3	1413	20	AAE63560	Islet cell antiod
278	15	5.3	690	8	AAE71375	Sequence of (2'-5'	351	15	5.3	1413	21	AAE46459	Pancreatic islet c
279	15	5.3	690	16	AAE00620	Human (2'-5') oli	352	15	5.3	1419	21	AAC55827	S. Lavendulae MmcM
280	15	5.3	700	22	AAH93172	Human inflammatory	353	15	5.3	1419	22	AAE80491	Cell proliferation
281	15	5.3	700	22	AAH93173	Human inflammatory	354	15	5.3	1439	21	AAC41159	Arabidopsis thalia
282	15	5.3	710	21	AAE12499	Aspergillus oryzae	355	15	5.3	1449	22	AAH17092	Human CDNA sequenc
283	15	5.3	710	21	AAC58891	Human tumour supp	356	15	5.3	1479	19	AAV34206	Human CDNA sequenc
284	15	5.3	720	22	AAH05025	Human CDNA clone (357	15	5.3	1479	20	AAAX04350	Human secreted pro
285	15	5.3	735	22	AAH33209	Human colon cancer	358	15	5.3	1479	22	AAI60623	Human polynucleoti
286	15	5.3	741	19	AAV59520	Human secreted pro	359	15	5.3	1480	22	AAH16004	Human CDNA sequenc
287	15	5.3	766	22	AAH01698	Human colon cancer	360	15	5.3	1485	22	AAE94451	Human hydrophobic
288	15	5.3	766	22	AAH03477	Human CDNA clone (361	15	5.3	1493	22	AAE58837	Human polynucleoti
289	15	5.3	780	13	AAO29859	Odorant receptor c	362	15	5.3	1496	22	AAE58205	Human angiotensino
290	15	5.3	788	22	AAH05650	Human CDNA clone (363	15	5.3	1500	21	AAE29584	Japanese encephali
291	15	5.3	792	22	AAH08239	Human CDNA clone (364	15	5.3	1501	22	AAE89854	Nucleotide sequenc
292	15	5.3	803	22	AAH03713	Human CDNA clone (365	15	5.3	1505	22	AAH13736	Human CDNA sequenc
293	15	5.3	826	22	AAH08004	Human CDNA clone (366	15	5.3	1527	21	AAH16633	Human secreted pro
294	15	5.3	837	21	AAE09000	Fusarium venenatum	367	15	5.3	1532	22	AAH15439	Human CDNA sequenc
295	15	5.3	852	22	AAE24905	Nucleotide sequenc	368	15	5.3	1537	19	AAE59804	Human secreted pro
296	15	5.3	854	19	AAE60001	Nucleic acid G255	369	15	5.3	1572	21	AAE66002	E. coli proliferat
297	15	5.3	854	22	AAE24906	Nucleotide sequenc	370	15	5.3	1590	16	AAE00622	Human (2'-5') olig
298	15	5.3	936	22	AAH05357	Human CDNA clone (371	15	5.3	1605	17	AAE59687	Human secreted pro
299	15	5.3	936	22	AAH98153	Human EST-derived	372	15	5.3	1607	17	AAE35277	Chemokine receptor
300	15	5.3	936	22	AAH70577	Human cervical can	373	15	5.3	1614	18	AAE84706	Human angiotensino
301	15	5.3	973	21	AAE41634	Arabidopsis thalia	374	15	5.3	1621	22	AAE98633	Human EST-derived
302	15	5.3	1025	21	AAD00813	Human irritable bo	375	15	5.3	1630	16	AAE03904	Human 2-5A synthe
303	15	5.3	1039	21	AAD00814	Human irritable bo	376	15	5.3	1650	18	AAE59645	Human 2-5A synthe

377	15	5.3	1661	22	AAAF81731	Human protease and	C 450	15	5.3	3341	22	AAH29722	S cerevisiae apopt
C 378	15	5.3	1677	21	AAAS8872	DNA encoding a hum	C 451	15	5.3	3349	18	AAAT51256	Human AD4 gene gen
379	15	5.3	1700	21	AAZ93706	CYB5RP fatty acid	C 452	15	5.3	3383	12	AAH16748	Human CDNA sequenc
380	15	5.3	1723	22	AAH15211	Human CDNA sequenc	C 453	15	5.3	3396	15	AAO55515	Prpase PRP35 genom
C 381	15	5.3	1752	22	AAAF57408	Human p95 protein	C 454	15	5.3	3503	19	AAV44041	mouse bRFG recepto
382	15	5.3	1757	21	AAAG0954	Human fatty acid d	C 455	15	5.3	3561	15	AAO55970	prpase PRP35 DNA s
C 383	15	5.3	1760	12	AAO13396	Microsomal dipepti	C 456	15	5.3	3613	20	AAH67627	Human islet cell a
C 384	15	5.3	1772	22	AAH99727	Human protein enco	C 457	15	5.3	3867	21	AAAC3304	Arabidopsis thalia
385	15	5.3	1803	22	AAAC84312	Human EXCS encodin	C 458	15	5.3	4011	22	AAAD06820	Human adenylate cy
386	15	5.3	1807	22	AAH34820	Human colon cancer	C 459	15	5.3	4024	22	AAH84595	MTG16b protein cod
387	15	5.3	1821	22	AAAS9865	Human polynucleoti	C 460	15	5.3	4080	19	AAV19480	Human TP33 CDNA.
C 388	15	5.3	1832	21	AAAF18039	Lung cancer associ	C 461	15	5.3	4080	20	AAAO1534	Human TP33 gene.
389	15	5.3	1860	22	AAAF99965	Human colon carcin	C 462	15	5.3	4226	20	AAH4594	MTG16a protein cod
C 390	15	5.3	1897	21	AAZ45803	Human adult skin c	C 463	15	5.3	4300	13	AAO25812	Clone 45-A. Synth
391	15	5.3	1899	22	AAAS02641	Human secreted pro	C 464	15	5.3	4579	20	AAAX20571	Polynucleotide seq
392	15	5.3	1900	10	AAAN91492	Escherichia coli m	C 465	15	5.3	4602	22	AAIS9330	Human polynucleoti
393	15	5.3	1918	22	AAH17464	Human CDNA sequenc	C 466	15	5.3	4670	22	AAI60734	Human polynucleoti
394	15	5.3	1971	22	AAAF45089	Human secreted pro	C 467	15	5.3	4758	22	AAI58948	Human polynucleoti
395	15	5.3	1989	21	AAO78111	Human cancer assoc	C 468	15	5.3	4799	21	AAAC76258	Human OREFX OREF1813
396	15	5.3	2005	15	AAO71303	Mouse osteoblast-s	C 469	15	5.3	5136	22	AAH44289	Human ribosome pro
397	15	5.3	2016	19	AAV34262	Human secreted pro	C 470	15	5.3	5154	20	AAAX22554	Human ZEGBP1 CDNA.
398	15	5.3	2043	22	AAH57373	Human heart cell s	C 471	15	5.3	5236	13	AAO25815	Clone 45-A. subtyp
399	15	5.3	2099	22	AAAC91605	Human angiotensino	C 472	15	5.3	5274	8	AAAT70679	Sequence encoding
400	15	5.3	2099	22	AAAC91678	Human angiotensino	C 473	15	5.3	5274	8	AAAT70824	Catechol producing
401	15	5.3	2099	22	AAAC91679	Human angiotensino	C 474	15	5.3	5449	21	AAAC64598	Streptococcus pneu
402	15	5.3	2099	22	AAAC91680	Human angiotensino	C 475	15	5.3	5555	21	AAAC77191	Human OREFX OREF2746
403	15	5.3	2099	22	AAAC91681	Human angiotensino	C 476	15	5.3	5596	20	AAAS9378	Aspergillus nidula
404	15	5.3	2099	22	AAAC91682	Human angiotensino	C 477	15	5.3	5640	22	AAI58155	Human polynucleoti
405	15	5.3	2099	22	AAAC91683	Human angiotensino	C 478	15	5.3	5828	22	AAH84592	Human polynucleoti
406	15	5.3	2099	22	AAAC91684	Human angiotensino	C 479	15	5.3	5943	22	AAAS2920	AML1-MTG16 fusion
407	15	5.3	2123	21	AAAC60997	Human vesicle asso	C 480	15	5.3	6056	22	AAH84593	DNA encoding novel
408	15	5.3	2171	10	AAAN90616	CPW32b CDNA. AAN	C 481	15	5.3	6701	22	AAH16661	AML1-MTG16 fusion
409	15	5.3	2171	13	AAO21175	Human CDW32b antiq	C 482	15	5.3	7812	22	AAAI2962	Human CDNA sequenc
410	15	5.3	2171	17	AAAT14715	Human CDW32b antiq	C 483	15	5.3	7827	22	AAAI58227	Enterococcus faeca
411	15	5.3	2171	19	AAV63453	Human CDW32b antiq	C 484	15	5.3	8011	19	AAV38336	Human polynucleoti
412	15	5.3	2171	20	AAV81211	Human CDW32b antiq	C 485	15	5.3	8065	19	AAV38335	Manic-depressive I
413	15	5.3	2171	21	AAAS0589	Human cell surface	C 486	15	5.3	8068	21	AAZ45340	Complete nucleotid
414	15	5.3	2171	22	AAAS05183	Human lymphocyte c	C 487	15	5.3	8238	15	AAAS9078	Nucleotide sequenc
415	15	5.3	2201	11	AAO06005	Sequence encoding	C 488	15	5.3	8535	15	AAO73731	GalV SEATO genome.
416	15	5.3	2282	22	AAI59175	Human polynucleoti	C 489	15	5.3	10427	21	AAZ36325	Mechanical stress
417	15	5.3	2293	20	AAZ11899	Human potassium ch	C 490	15	5.3	11570	21	AAAS95905	Human RIK-L2 gene.
418	15	5.3	2345	22	AAH16123	Human CDNA sequenc	C 491	15	5.3	13114	21	AAAF60744	Rice gene for resi
C 419	15	5.3	2345	20	AAH18951	Human P10R-1 encod	C 492	15	5.3	15692	20	AAAX24731	London-FAD App tar
420	15	5.3	2353	14	AAO37948	Sequence of a DNA	C 493	15	5.3	15692	20	AAAX24732	Swedish/London-FAD
421	15	5.3	2353	19	AAV07530	Human transcrip tio	C 494	15	5.3	15701	20	AAAX24733	Swedish-FAD APP713
422	15	5.3	2353	22	AAH02898	Human shear stress	C 495	15	5.3	18402	21	AAZ93705	CYB5RP fatty acid
423	15	5.3	2367	21	AAZ92178	N-acetylglucosamin	C 496	15	5.3	35100	20	AAV73803	KSHV IUR DNA (nucl
C 424	15	5.3	2386	22	AAAC66918	Rat G-protein modu	C 497	15	5.3	45265	20	AAZ46508	Sequence of a COSM
425	15	5.3	2390	22	AAH14062	Human CDNA sequenc	C 498	15	5.3	45624	22	AAE88315	S. spinosa DNA fra
426	15	5.3	2403	21	AAAB8288	Adenovirus thymidi	C 499	15	5.3	50000	22	AAE88312	S. spinosa DNA fra
427	15	5.3	2421	21	AAZ36409	CDNA encoding a hu	C 500	15	5.3	53500	21	AAAC55842	Complete nucleotid
428	15	5.3	2452	22	AAH17574	Human CDNA sequenc	C 501	15	5.3	80161	20	AAZ21501	DNA fragment of Sa
429	15	5.3	2460	22	AAH16954	Human CDNA sequenc	C 502	15	5.3	95223	21	AAAF22282	BAC containing rep
C 430	15	5.3	2503	22	AAH14503	Human CDNA sequenc	C 503	15	5.3	109519	22	AAAS08693	Micromonospora DNA
431	15	5.3	2517	21	AAZ51620	Human membrane cha	C 504	15	5.3	133719	21	AAAC64754	Macaca mulatta rha
432	15	5.3	2575	19	AAZ98572	DNA encoding a S.	C 505	15	5.3	160552	22	AAAD02697	Human glycosyl sul
433	15	5.3	2624	15	AAO54846	Sequence encoding	C 506	15	5.3	305107	22	AAH62689	Shrimp white spot
434	15	5.3	2624	21	AAZ92177	N-acetylglucosamin	C 507	15	5.3	309400	22	AAH68534	C glutamicum codin
435	15	5.3	2650	21	AAZ93150	Sequence encoding	C 508	15	5.3	910715	22	AAAT30248	Borrelia burgdorfe
436	15	5.3	2751	21	AAAF16253	Human prostate can	C 509	15	5.3	22	17	AAAT35281	ChemoKine receptor
437	15	5.3	2770	21	AAAC93410	Human secreted pro	C 510	15	5.0	22	20	AAAX35968	Primer used to amp
C 438	15	5.3	2825	19	AAZ96282	S. pneumoniae deri	C 511	15	5.0	23	20	AAAZ2310	Mouse NAKC protein
439	15	5.3	2892	22	AAH17602	Human CDNA sequenc	C 512	15	5.0	25	22	AAH27859	5' RACE primer use
C 440	15	5.3	2937	22	AAH18341	Human CDNA sequenc	C 513	15	5.0	26	21	AAAS9558	PCR primer used to
441	15	5.3	2991	22	AAH16626	Human CDNA sequenc	C 514	15	5.0	31	22	AAI31138	Human single nucle
C 442	15	5.3	3113	19	AAV46322	Human secreted pro	C 515	15	5.0	41	22	AAAS01735	Glucanase gene mut
C 443	15	5.3	3113	22	AAAF9489	Human CDNA clone E	C 516	15	5.0	51	21	AAAT7378	Human clone cg4491
444	15	5.3	3236	19	AAV52313	Streptococcus pneu	C 517	15	5.0	51	21	AAAT7379	Human clone cg4491
445	15	5.3	3294	22	AAH73225	Human cervical can	C 518	15	5.0	51	22	AAH90417	Human clone cg4393
C 446	15	5.3	3311	20	AAZ32342	Human pancreatic i	C 519	15	5.0	51	22	AAH90418	Human clone cg4393
C 447	15	5.3	3311	20	AAZ09923	Human islet cell a	C 520	15	5.0	52	20	AAAX78261	TAT-HV-P15B1d fusi
C 448	15	5.3	3311	20	AAV63561	Islet cell anti bod	C 521	15	5.0	52	21	AAAC63847	Mouse B1d p15 doma
C 449	15	5.3	3311	21	AAZ46460	Pancreatic islet c	C 522	15	5.0	52	21	AAA46601	PCR primer used to

523	14	5.0	52	22	AAH21149	CD95L construct DN	596	14	5.0	370	22	AAC9641	Skin cell cDNA, SE
524	14	5.0	56	22	AAH91973	Human inflammatory	597	14	5.0	371	22	AAI11991	Probe #1924 for ge
525	14	5.0	74	22	AAH74357	Nucleotide sequenc	598	14	5.0	371	22	AAI33323	Probe #2009 used t
526	14	5.0	83	22	AAI54139	Probe #22825 used	599	14	5.0	371	22	AAI01923	Probe #1914 used t
527	14	5.0	88	20	AAI78263	TAT-HV-p15Bid fusi	600	14	5.0	374	18	AAV78614	Staphylococcus aur
528	14	5.0	88	21	AAI63849	Mouse Bid p15 doma	601	14	5.0	375	19	AAV54056	DNA fragment 2 of
529	14	5.0	88	21	AAI46603	PCR primer used to	602	14	5.0	375	19	AAV54057	DNA fragment 3 of
530	14	5.0	97	22	AAI48010	Probe #16696 used	603	14	5.0	375	19	AAV54058	DNA fragment 4 of
531	14	5.0	98	21	AAI22730	Human secreted pro	604	14	5.0	377	22	AAI65917	Novel human polyu
532	14	5.0	102	16	AAO81963	Human CD44 variabl	605	14	5.0	379	21	AAH30724	Human colon cancer
533	14	5.0	111	22	AAI09764	Probe #9755 used t	606	14	5.0	379	21	AAH31106	Human colon cancer
534	14	5.0	114	16	AAI24782	Human gene signatu	607	14	5.0	380	22	AAI64452	Novel human polyu
535	14	5.0	149	22	AAI54758	Probe #23444 used	608	14	5.0	381	22	AAI64655	Novel human polyu
536	14	5.0	157	21	AAI21932	Human secreted pro	609	14	5.0	382	21	AAI61567	Genetic suppressor
537	14	5.0	165	17	AAI13568	Oligo 1 for constr	610	14	5.0	382	22	AAI27676	Probe #17609 for g
538	14	5.0	165	18	AAI73983	Rat THY-1 synthe	611	14	5.0	386	20	AAI51920	Human secreted pro
539	14	5.0	165	19	AAI23321	Rat THY-1 synthe	612	14	5.0	386	21	AAI42676	Human 5' EST isola
540	14	5.0	166	16	AAI19688	Human gene signatu	613	14	5.0	387	22	AAI04780	Probe #4771 used t
541	14	5.0	166	16	AAI19688	Human gene signatu	614	14	5.0	388	22	AAI66902	Novel human polyu
542	14	5.0	169	21	AAI22524	Human gene signatu	615	14	5.0	389	21	AAH30725	Human colon cancer
543	14	5.0	170	16	AAI19683	Human gene signatu	616	14	5.0	389	21	AAI60947	Human colon cancer
544	14	5.0	171	21	AAI45138	Human secreted exp	617	14	5.0	390	22	AAI60947	Human polynucleoti
545	14	5.0	172	20	AAV90378	EST clone DM169.	618	14	5.0	390	22	AAI60947	Novel human polyu
546	14	5.0	179	22	AAI24046	Probe #13979 for g	619	14	5.0	392	22	AAI66296	Novel human polyu
547	14	5.0	179	22	AAI49347	Probe #18033 used	620	14	5.0	401	21	AAI77128	Novel human polyu
548	14	5.0	191	22	AAI09630	Probe #9621 used t	621	14	5.0	401	21	AAI77128	Novel human polyu
549	14	5.0	191	21	AAI09630	Human secreted pro	622	14	5.0	403	21	AAI69888	Human ovarian carc
550	14	5.0	193	22	AAI50359	Probe #19045 used	623	14	5.0	404	22	AAI19559	Human secreted exp
551	14	5.0	208	22	AAI51475	Probe #20161 used	624	14	5.0	404	22	AAI44749	Probe #13435 used
552	14	5.0	211	19	AAV18436	Partial ribosomal	625	14	5.0	404	22	AAI05280	Probe #5271 used t
553	14	5.0	212	22	AAI72844	Gene #27 associate	626	14	5.0	405	14	AAQ39779	Expressed Sequence
554	14	5.0	218	21	AAI29997	Human secreted pro	627	14	5.0	405	14	AAQ39779	Human brain Expres
555	14	5.0	222	20	AAV05181	Human MSH5 (hMSH5)	628	14	5.0	405	21	AAH52183	Human AFP protein
556	14	5.0	225	20	AAV88217	EST clone GG440.	629	14	5.0	406	21	AAA43888	Human secreted exp
557	14	5.0	240	21	AAI42686	Human 5' EST isola	630	14	5.0	407	21	AAH30312	Human colon cancer
558	14	5.0	242	21	AAI98009	Human T gene DNA f	631	14	5.0	407	22	AAI67032	Novel human polyu
559	14	5.0	246	18	AAV78973	Staphylococcus aur	632	14	5.0	410	21	AAI64963	Membrane-bound pro
560	14	5.0	249	14	AAO42774	Ligand-induced gen	633	14	5.0	410	22	AAI64963	Human EST DNA3906
561	14	5.0	252	22	AAO09458	Human aspartyl pro	634	14	5.0	412	21	AAI27757	Human secreted pro
562	14	5.0	254	20	AAI27261	Prostate-tumour de	635	14	5.0	419	21	AAH30406	Human colon cancer
563	14	5.0	251	21	AAI18503	Human secreted pro	636	14	5.0	419	21	AAI007447	Human secreted pro
564	14	5.0	264	22	AAI20207	Probe #10140 for g	637	14	5.0	422	18	AAI88808	Nuclear steroid ho
565	14	5.0	264	22	AAI45407	Probe #14093 used	638	14	5.0	423	21	AAI77723	CDNA encoding huma
566	14	5.0	264	22	AAI05911	Probe #5902 used t	639	14	5.0	423	22	AAI28461	Colon tumour relat
567	14	5.0	266	21	AAI1895	Human secreted pro	640	14	5.0	424	21	AAI02693	Human secreted pro
568	14	5.0	272	16	AAI22195	Human gene signatu	641	14	5.0	425	21	AAI22354	Human secreted pro
569	14	5.0	279	21	AAI00392	Human colon cancer	642	14	5.0	430	22	AAI14776	Probe #4709 for ge
570	14	5.0	279	22	AAI28148	Probe #18081 for g	643	14	5.0	430	22	AAI36134	Probe #4820 used t
571	14	5.0	281	16	AAI57185	Probe #25871 used	644	14	5.0	432	22	AAI04569	Probe #4560 used t
572	14	5.0	281	16	AAI22991	Human gene signatu	645	14	5.0	432	22	AAI92325	Bovine mammary tis
573	14	5.0	294	14	AAO61096	Human brain Expres	646	14	5.0	433	22	AAI14857	Probe #4790 for ge
574	14	5.0	300	20	AAI13914	Human gene expres	647	14	5.0	433	22	AAI36217	Probe #4903 used t
575	14	5.0	300	21	AAI00584	Human colon cancer	648	14	5.0	433	22	AAI04645	Probe #4636 used t
576	14	5.0	303	20	AAV88058	EST clone FS10. H	649	14	5.0	435	21	AAI04645	Human secreted pro
577	14	5.0	312	21	AAI26275	Human secreted pro	650	14	5.0	435	22	AAI74376	Nucleotide sequenc
578	14	5.0	316	21	AAI00710	Human secreted pro	651	14	5.0	441	21	AAI07234	Human secreted pro
579	14	5.0	320	20	AAV86549	EST clone A2191.	652	14	5.0	444	22	AAI21195	Probe #11128 for g
580	14	5.0	321	21	AAI02402	Human secreted pro	653	14	5.0	444	22	AAI46463	Probe #15149 used
581	14	5.0	322	21	AAI04385	Human secreted pro	654	14	5.0	444	22	AAI06909	Probe #6900 used t
582	14	5.0	329	21	AAI27249	Human secreted pro	655	14	5.0	454	22	AAI10972	Probe #905 for gen
583	14	5.0	333	14	AAO59791	Human brain Expres	656	14	5.0	454	22	AAI32232	Probe #918 used to
584	14	5.0	336	21	AAI04127	Human secreted pro	657	14	5.0	454	22	AAI00893	Probe #884 used to
585	14	5.0	338	21	AAI41587	Human secreted pro	658	14	5.0	457	22	AAI11852	Probe #1785 for ge
586	14	5.0	340	21	AAI77781	CDNA encoding huma	659	14	5.0	457	22	AAI33173	Probe #1859 used t
587	14	5.0	340	22	AAI28519	Colon tumour relat	660	14	5.0	457	22	AAI01789	Probe #1780 used t
588	14	5.0	350	21	AAI61571	Genetic suppressor	661	14	5.0	460	18	AAI73391	S. pneumoniae HSP7
589	14	5.0	352	14	AAO60686	Human brain Expres	662	14	5.0	464	22	AAI16007	Probe #5940 for ge
590	14	5.0	354	20	AAI27018	Differentially exp	663	14	5.0	464	22	AAI38061	Probe #6747 used t
591	14	5.0	354	21	AAI15760	Human secreted pro	664	14	5.0	468	21	AAI098027	Human colon cancer
592	14	5.0	355	22	AAH83480	Human ovarian tumo	665	14	5.0	468	22	AAI15808	Probe #5741 for ge
593	14	5.0	358	21	AAI61704	CDNA encoding mur	666	14	5.0	468	22	AAI37169	Probe #5855 used t
594	14	5.0	358	22	AAI09637	Skin cell cDNA, SE	667	14	5.0	468	22	AAI37618	Probe #6304 used t
595	14	5.0	370	21	AAI61708	CDNA encoding mur	668	14	5.0	471	22	AAI34845	Probe #3531 used t

C 669	14	5.0	474	21	AAC00717	Human secreted pro	742	14	5.0	601	21	AAF08462	Fusarium venenatum
C 670	14	5.0	474	22	AAI59933	Human polynucleoti	C 743	14	5.0	622	21	AAFI5986	Human prostate can
C 671	14	5.0	475	22	AAI38339	Probe #7025 used t	C 744	14	5.0	634	20	AAZ77563	Human ovarian tumo
C 672	14	5.0	479	22	AAI11310	Probe #1243 for ge	C 745	14	5.0	634	20	AAZ41361	Human normal uteru
C 673	14	5.0	479	22	AAI32578	Probe #1264 used t	C 746	14	5.0	636	21	AAC39050	Human pancreatic c
C 674	14	5.0	479	22	AAI01225	Probe #1216 used t	C 747	14	5.0	639	20	AAC33389	Human Delta7-stero
C 675	14	5.0	482	20	AAI20882	Polynucleotide seq	C 748	14	5.0	645	21	AAFI4048	Aspergillus oryzae
C 676	14	5.0	484	22	AAI32239	Mouse urinary tryp	C 749	14	5.0	649	21	AAI25041	Corn transcription
C 677	14	5.0	485	22	AAI15126	Probe #5059 for ge	C 750	14	5.0	652	20	AAI22636	MDA-231 PBR (perip
C 678	14	5.0	486	22	AAH66643	C glutamicum codin	C 751	14	5.0	652	20	AAI22637	MCF-7 PBR (periphe
C 679	14	5.0	489	22	AAH83600	Human ovarian tumo	C 752	14	5.0	655	21	AAAI6335	Human colon cancer
C 680	14	5.0	491	22	AAI10757	Probe #690 for gen	C 753	14	5.0	656	20	AAAI8337	Mouse IMC carcinom
C 681	14	5.0	491	22	AAI32015	Probe #701 used to	C 754	14	5.0	669	20	AAK61367	DNA encoding a hum
C 682	14	5.0	491	22	AAI00684	Probe #675 used to	C 755	14	5.0	675	21	AAFI4000	Aspergillus oryzae
C 683	14	5.0	497	16	AAI19490	Human gene signatu	C 756	14	5.0	675	21	AAFI67810	Corynebacterium g1
C 684	14	5.0	497	21	AAI77569	Human ORFX ORF3124	C 757	14	5.0	677	21	AAFI3021	Aspergillus oryzae
C 685	14	5.0	500	21	AAI70362	Human peripheral-t	C 758	14	5.0	678	21	AAFI3659	Aspergillus oryzae
C 686	14	5.0	500	21	AAI72047	Human PTBR CDNA f	C 759	14	5.0	685	22	AAH53236	S. epidermidis ope
C 687	14	5.0	504	22	AAI41723	Probe #10409 used	C 760	14	5.0	685	22	AAH07876	Human CDNA clone (
C 688	14	5.0	510	21	AAC00714	Human secreted pro	C 761	14	5.0	688	20	AAV86183	EST clone J546. H
C 689	14	5.0	515	21	AAC03789	Probe #7727 for ge	C 762	14	5.0	688	22	AAH34559	Human colon cancer
C 690	14	5.0	515	22	AAI17794	Probe #11442 used	C 763	14	5.0	692	21	AAFI2562	Aspergillus oryzae
C 691	14	5.0	515	22	AAI42756	Human TAMGO 273 OR	C 764	14	5.0	695	22	AAH04659	Human CDNA clone (
C 692	14	5.0	516	22	AAI45128	Arabidopsis thalia	C 765	14	5.0	700	22	AAH92829	Human inflammatory
C 693	14	5.0	518	22	AAC54306	Probe #9785 used t	C 766	14	5.0	700	22	AAH92943	Human inflammatory
C 694	14	5.0	518	22	AAH13614	Human CDNA clone (C 767	14	5.0	703	22	AAFI1563	Corynebacterium g1
C 695	14	5.0	519	22	AAH13614	Human CDNA clone (C 768	14	5.0	705	21	AAC98085	Human colon cancer
C 696	14	5.0	528	20	AAI233651	Human breast tumou	C 769	14	5.0	709	21	AAC64439	Lycopodium escul
C 697	14	5.0	530	20	AAH06271	Human CDNA clone (C 770	14	5.0	709	21	AAC64449	Lycopodium escul
C 698	14	5.0	538	21	AAFI2510	Aspergillus oryzae	C 771	14	5.0	711	22	AAH04777	Human CDNA clone (
C 699	14	5.0	544	22	AAH10679	Human CDNA clone (C 772	14	5.0	713	21	AAFI1415	Aspergillus oryzae
C 700	14	5.0	550	21	AAI70361	Rat peripheral-tyr	C 773	14	5.0	714	22	AAI28682	Human protein Hp03
C 701	14	5.0	550	21	AAH72046	Rat PTBR CDNA frag	C 774	14	5.0	715	21	AAFI2419	Aspergillus oryzae
C 702	14	5.0	552	21	AAH82610	Human ovarian tumo	C 775	14	5.0	715	21	AAI98478	Bovine PTEN/MMAC1
C 703	14	5.0	554	22	AAI41642	Probe #10328 used	C 776	14	5.0	717	22	AAH04962	Human CDNA clone (
C 704	14	5.0	558	10	AAH90386	Fish growth hormon	C 777	14	5.0	719	21	AAFI1997	Aspergillus oryzae
C 705	14	5.0	558	22	AAH09096	Human CDNA clone (C 778	14	5.0	722	21	AAC46860	Arabidopsis thalia
C 706	14	5.0	561	22	AAI18315	Probe #8248 for ge	C 779	14	5.0	722	22	AAH07074	Human CDNA clone (
C 707	14	5.0	561	22	AAI43390	Probe #12076 used	C 780	14	5.0	726	22	AAH03817	Human protein enco
C 708	14	5.0	562	22	AAH82974	Human ovarian tumo	C 781	14	5.0	728	22	AAH99218	Fusarium venenatum
C 709	14	5.0	562	22	AAH06252	Human CDNA clone (C 782	14	5.0	729	21	AAFI09999	Aspergillus oryzae
C 710	14	5.0	567	22	AAH12207	Human CDNA clone (C 783	14	5.0	731	21	AAFI2405	Aspergillus oryzae
C 711	14	5.0	568	22	AAI41969	Probe #10655 used	C 784	14	5.0	737	21	AAI67075	Euclalyptus grandis
C 712	14	5.0	569	15	AAO62833	GGF gene coding se	C 785	14	5.0	739	14	AAO45082	Human glucokinase
C 713	14	5.0	569	15	AAO58316	GGF gene coding se	C 786	14	5.0	741	22	AAI03835	Human carboxypepti
C 714	14	5.0	569	16	AAO74905	Bovine glial cell	C 787	14	5.0	742	18	AAV77973	Staphylococcus aur
C 715	14	5.0	569	17	AAI48074	Bovine glial growt	C 788	14	5.0	743	20	AAI98681	Human validated ca
C 716	14	5.0	569	17	AAI30983	Bovine glial growt	C 789	14	5.0	747	21	AAI01873	Human colon cancer
C 717	14	5.0	569	17	AAI06724	Bovine glial growt	C 790	14	5.0	748	21	AAI88937	Human validated ca
C 718	14	5.0	569	20	AAI81193	Bovine neuregulin	C 791	14	5.0	749	21	AAI09317	Human cancer assoc
C 719	14	5.0	569	22	AAI19005	Probe #8938 for ge	C 792	14	5.0	750	20	AAI215616	Human gene express
C 720	14	5.0	569	22	AAI44143	Probe #12829 used	C 793	14	5.0	752	20	AAI215058	Human gene express
C 721	14	5.0	570	22	AAH83876	Human ovarian tumo	C 794	14	5.0	757	22	AAH07989	Human CDNA clone (
C 722	14	5.0	572	22	AAH83548	Human ovarian tumo	C 795	14	5.0	761	17	AAI37365	Rat phospholipid h
C 723	14	5.0	577	19	AAV34220	Human secreted pro	C 796	14	5.0	767	21	AAI81592	N. meningitidis pa
C 724	14	5.0	578	22	AAI18679	Probe #6612 for ge	C 797	14	5.0	770	21	AAI67146	Euclalyptus grandis
C 725	14	5.0	579	22	AAH52280	Human AFP protein	C 798	14	5.0	771	22	AAI10018	Blocker molecule #
C 726	14	5.0	580	22	AAI18158	Probe #8091 for ge	C 799	14	5.0	773	22	AAH11540	Human cervical can
C 727	14	5.0	580	22	AAI43167	Probe #11853 used	C 800	14	5.0	774	21	AAC70797	Single nucleotide
C 728	14	5.0	580	22	AAI02713	Human epididymal p	C 801	14	5.0	774	21	AAC70857	Single nucleotide
C 729	14	5.0	583	12	AAO13127	HE4 epididymal-spe	C 802	14	5.0	775	22	AAH74327	Nucleotide sequenc
C 730	14	5.0	583	19	AAV64622	Human H24 CDNA. H	C 803	14	5.0	776	20	AAI15918	Human gene express
C 731	14	5.0	583	21	AAC53813	Human epididymal p	C 804	14	5.0	776	20	AAI98785	Human validated ca
C 732	14	5.0	583	22	AAI02714	Human epididymal p	C 805	14	5.0	779	19	AAI26428	S. pneumoniae deri
C 733	14	5.0	584	21	AAC58371	Human PRO619 nucle	C 806	14	5.0	779	22	AAI89263	Rat cannelito 4 (Rom
C 734	14	5.0	584	21	AAI264983	Membrane-bound pro	C 807	14	5.0	780	22	AAH05021	Human CDNA clone (
C 735	14	5.0	584	22	AAI21444	Human cDNA sequenc	C 808	14	5.0	782	22	AAH04542	Human CDNA clone (
C 736	14	5.0	584	22	AAI4129	Human PRO619 (UNQ3	C 809	14	5.0	784	21	AAC95091	Human pancreatic c
C 737	14	5.0	585	22	AAI40962	Probe #9648 used t	C 810	14	5.0	784	22	AAH05499	Human CDNA clone (
C 738	14	5.0	586	21	AAC74925	Human ORFX ORF480	C 811	14	5.0	785	21	AAI46666	CDNA of a human pe
C 739	14	5.0	588	19	AAV22147	Mouse BH3 interact	C 812	14	5.0	789	18	AAI79326	DNA encoding Therm
C 740	14	5.0	598	20	AAV22147	Human cDNA clone (C 813	14	5.0	792	21	AAFI09601	Fusarium venenatum
C 741	14	5.0	600	20	AAV88046	EST clone FP268.	C 814	14	5.0	792	22	AAI24306	Probe #14239 For g

815	14	5	0	794	22	AA05263	Human secreted pro	C	888	14	5	0	1068	22	AAH81443	Escherichia coli p
816	14	5	0	798	22	AA23513	DNA encoding prote	C	889	14	5	0	1080	22	AAH9256	Mouse camello 2 (M
817	14	5	0	802	21	AA02132	Human colon cancer	C	890	14	5	0	1085	22	AAH64893	Human secreted pro
818	14	5	0	807	22	AAH52244	Human AFP protein	C	891	14	5	0	1086	11	AAQ03971	Gene encoding D-am
819	14	5	0	821	21	AAH70363	Human peripheral-t	C	892	14	5	0	1100	22	AAQ02050	DNA encoding molec
820	14	5	0	821	21	AAH72048	DNA encoding human	C	893	14	5	0	1101	22	AA160897	Human polynucleoti
821	14	5	0	821	21	AAH46678	CDNA of a human pe	C	894	14	5	0	1104	21	AD01956	Murine TH2AF1 isot
822	14	5	0	827	21	AA259143	Human Tumour metas	C	895	14	5	0	1104	21	AD01957	Murine TH2AF1 isot
C 823	14	5	0	828	22	AAH07393	Human CDNA clone (C	896	14	5	0	1104	21	AD01959	Murine TH2AF1 isot
C 824	14	5	0	830	22	AAH07946	Human CDNA clone (C	897	14	5	0	1104	21	AD01960	Murine TH2AF1 isot
C 825	14	5	0	831	22	AAH04643	Human CDNA clone (C	898	14	5	0	1126	21	AA47645	Arabidopsis thalia
826	14	5	0	831	22	AAH05352	Human CDNA clone (C	899	14	5	0	1156	22	AD11066	Human secreted pro
827	14	5	0	833	22	AAH08273	Human CDNA clone (C	900	14	5	0	1160	19	AAV62135	HSV-2 strain SB5 C
C 828	14	5	0	833	22	AA160632	Human polynucleoti	C	901	14	5	0	1167	22	AA161143	Human polynucleoti
C 829	14	5	0	854	22	AA158846	Human polynucleoti	C	902	14	5	0	1169	22	AA159357	Human polynucleoti
C 830	14	5	0	855	21	AA00718	Human secreted pro	C	903	14	5	0	1177	22	AAH34085	Human colon cancer
831	14	5	0	858	20	AA06225	Human secreted pro	C	904	14	5	0	1184	21	AA36340	Arabidopsis thalia
832	14	5	0	858	22	AAH03335	Human CDNA clone (C	905	14	5	0	1184	21	AA50834	Arabidopsis thalia
C 833	14	5	0	859	22	AAH05092	Human CDNA clone (C	906	14	5	0	1187	21	AACT5882	Human ORFX ORF147
C 834	14	5	0	860	22	AAH95507	Human protein enco	C	907	14	5	0	1191	22	AAH74326	Nucleotide sequenc
C 835	14	5	0	861	7	AAH60938	Equine IFN-alpha-4	C	908	14	5	0	1192	22	AAH72825	Secreted protein g
C 836	14	5	0	871	21	AA259142	Metastasis-associat	C	909	14	5	0	1204	8	AAH70816	Aspergillus nidula
C 837	14	5	0	882	20	AA28026	Adipogenesis inhib	C	910	14	5	0	1212	16	AAH09086	Fusarium heterospo
C 838	14	5	0	882	21	AA096320	Human adipocytogen	C	911	14	5	0	1221	22	AD08436	Human secreted pro
C 839	14	5	0	888	21	AA098016	Human colon cancer	C	912	14	5	0	1222	21	AA258189	Arabidopsis phosph
C 840	14	5	0	890	20	AA23642	Human breast tumor	C	913	14	5	0	1225	22	AAH27866	Nucleotide sequenc
C 841	14	5	0	891	21	AA00715	Human secreted pro	C	914	14	5	0	1235	20	AAH59630	CDNA encoding a hu
C 842	14	5	0	892	17	AA18979	Stannicalocalin alph	C	915	14	5	0	1235	22	AA159374	Human polynucleoti
C 843	14	5	0	892	21	AA239303	Human stannicalocal	C	916	14	5	0	1235	22	AD08064	Human extracellular
844	14	5	0	900	22	AA292506	Human T2R05 nucleot	C	917	14	5	0	1236	21	AACT6232	Human ORFX ORF1787
845	14	5	0	900	22	AAH92861	Rat camello 2 (Rcm	C	918	14	5	0	1242	21	AA36756	Human ORFX ORF1787
846	14	5	0	903	22	AAH89258	Mouse camello 4 (M	C	919	14	5	0	1243	12	AAQ12579	Coding sequence of
C 847	14	5	0	906	21	AA257116	Human stannicalocal	C	920	14	5	0	1257	17	AAH38266	D-epsilon oxidase ge
C 848	14	5	0	906	22	AA25370	Nucleotide sequenc	C	921	14	5	0	1257	21	AAH51899	54 gene different
C 849	14	5	0	908	22	AAH31030	Human stannicalocal	C	922	14	5	0	1257	22	AAH82610	Murine T helper ce
C 850	14	5	0	908	22	AAH98885	C elegans EST-deri	C	923	14	5	0	1257	22	AAH90887	Murine TH1 specific
C 851	14	5	0	909	22	AAH25831	Human adipogenesis	C	924	14	5	0	1257	22	AAH23459	Mouse 54 gene nuclei
852	14	5	0	911	21	AACT7565	Human ORFX ORF3120	C	925	14	5	0	1260	22	AAH28692	54 gene. Unidenti
C 853	14	5	0	913	21	AAAC4183	Arabidopsis thalia	C	926	14	5	0	1278	21	AAH15677	Human protein HP03
854	14	5	0	920	21	AACT7659	Human cancer assoc	C	927	14	5	0	1278	22	AAH13033	Human Asp2 nucleot
855	14	5	0	921	10	AAH90384	Fish growth hormon	C	928	14	5	0	1278	22	AAH11529	T7-Caspase-Caspase
856	14	5	0	921	11	AA006090	Sequence encoding	C	929	14	5	0	1278	22	AAH11714	T7-Caspase-Caspase
857	14	5	0	924	16	AAH03498	Human Fas ligand c	C	930	14	5	0	1278	22	AD06751	DNA encoding T7-ca
C 858	14	5	0	930	21	AA404938	Arabidopsis thalia	C	931	14	5	0	1284	20	AA25583	Human Rhog nucleot
C 859	14	5	0	932	22	AA500799	F. venenatum glucan	C	932	14	5	0	1284	22	AAH94726	Human CDNA encodin
C 860	14	5	0	936	20	AAH15651	Protein phosphatas	C	933	14	5	0	1286	22	AAH16346	Human CDNA sequenc
C 861	14	5	0	941	21	AACT74229	Human secreted pro	C	934	14	5	0	1287	21	AAH54072	Arabidopsis thalia
C 862	14	5	0	942	21	AACT9994	Human secreted pro	C	935	14	5	0	1287	22	AD13276	Human Asp2(b) deit
863	14	5	0	949	21	AACT7387	Arabidopsis thalia	C	936	14	5	0	1287	22	AAH11547	Human CDNA encodin
C 864	14	5	0	951	21	AACT6669	Human ORFX ORF224	C	937	14	5	0	1287	22	AAH11732	DNA encoding human
C 865	14	5	0	952	21	AA09063	Human pancreatic c	C	938	14	5	0	1287	22	AD06768	Human aspartyl pro
866	14	5	0	954	21	AA050190	Arabidopsis thalia	C	939	14	5	0	1298	11	AAQ02420	Pisamid PCSF3 cont
867	14	5	0	956	21	AA039796	Arabidopsis thalia	C	940	14	5	0	1298	21	AAH51618	Human membrane cta
868	14	5	0	972	19	AAV42194	Nucleotide sequenc	C	941	14	5	0	1302	21	AAH38476	Arabidopsis thalia
869	14	5	0	972	20	AAH35963	Nucleotide sequenc	C	942	14	5	0	1302	21	AAH15670	Human (pro-Asp-2(a)
870	14	5	0	972	21	AAH37898	Arabidopsis thalia	C	943	14	5	0	1302	22	AAH13032	Human (pro-Asp-2(a)
871	14	5	0	972	22	AA508454	Human CDNA encodin	C	944	14	5	0	1302	22	AAH11528	Human CDNA encodin
C 872	14	5	0	972	22	AA508459	Human Fas li	C	945	14	5	0	1302	22	AAH11713	DNA encoding T7-pu
C 873	14	5	0	975	21	AA49798	Arabidopsis thalia	C	946	14	5	0	1302	22	AD06750	Human (pro-Asp-2(a)
874	14	5	0	983	20	AAH18513	Mouse IMC carcino	C	947	14	5	0	1305	22	AD13277	Human Asp2(b) deit
875	14	5	0	990	21	AA289455	Mouse 3Gnt cDNA.	C	948	14	5	0	1305	22	AAH11548	Human CDNA encodin
C 876	14	5	0	1012	14	AAH90920	AT-1 gene and 3'no	C	949	14	5	0	1305	22	AAH11733	DNA encoding human
C 877	14	5	0	1014	14	AAQ04098	DNA encoding VCD4	C	950	14	5	0	1305	22	AD06769	Human aspartyl pro
C 878	14	5	0	1023	22	AAH65001	C glutamicum codin	C	951	14	5	0	1306	21	AAH13340	Aspergillus oryzae
C 879	14	5	0	1026	22	AAH55807	Human SCN2A genom	C	952	14	5	0	1308	21	AAH99092	Human ATP coding
C 880	14	5	0	1027	22	AA050600	Human DNA modificta	C	953	14	5	0	1308	22	AAH67352	C glutamicum codin
881	14	5	0	1027	22	AAH89255	Mouse camello 1 (M	C	954	14	5	0	1311	22	AD03837	Human carboxypepti
C 882	14	5	0	1041	21	AAH38035	Arabidopsis thalia	C	955	14	5	0	1315	21	AAH32752	Arabidopsis thalia
C 883	14	5	0	1047	22	AAH83992	Human interleukin-	C	956	14	5	0	1317	21	AAH47339	Arabidopsis thalia
C 884	14	5	0	1050	22	AD03838	Human carboxypepti	C	957	14	5	0	1319	21	AAH34872	Arabidopsis thalia
C 885	14	5	0	1065	21	AAH13118	Aspergillus oryzae	C	958	14	5	0	1332	21	AAH21724	Human breast and o
C 886	14	5	0	1066	20	AAH82169	P. vivax Pvs28 pol	C	959	14	5	0	1334	21	AAH95484	Human secreted pro
887	14	5	0	1066	22	AAH99436	Human protein enco	C	960	14	5	0	1334	21	AAH36785	Nucleotide sequenc

T7 Caspase-human P
T7-Human P10-Asp22
Human cDNA encoding
DNA encoding T7-hu
T7-Human-p10-Asp-2
Human EST-derived
Modified human asp
Human-Asp2(a) del
Human cDNA encoding
DNA encoding human
Human-Asp-2(a) del
pELC4 coding sequen
Nucleotide sequenc
Human GTP-binding
Polynucleotide seque
T7-caspase-human-p
Modified human asp
T7-Caspase-Human-P
Human-Asp2(a) del
Human T7-Caspase-h
Human cDNA encodin
DNA encoding T7-ca
DNA encoding human
T7-Caspase-human-P
Human-Asp-2(a) del
Human-Asp-2(a) del
Human stomach cell
Nucleotide sequenc
Human polynucleoti
Human cDNA sequenc
Human polynucleoti
Bacillus sp. metun
Nucleotide sequenc
Human SH3-containi
Arabidopsis thaliana
Arabidopsis thaliana
G-protein modulat
Human foetal cDNA,
lung cancer associ
Human cervical can

ALIGNMENTS

RESULT	1
AAFS5866	
ID	AAFS5866 standard; cDNA; 2087 BP
XX	
AC	
XX	AAFS5866;
XX	
DT	17-APR-2001 (first entry)
XX	
DE	Rat GLUTX1 coding sequence.
XX	
KW	Rat; GLUTX; gene therapy; vaccine
KW	hypoglycemia; glucose metabolism
KW	hypoglycemia; glucose metabolism
OS	Rattus sp.
XX	
PN	W0200104145.A2.
XX	
PD	18-JAN-2001.
XX	
PE	14-JUL-2000; 2000MO-IB01042.
XX	
PR	14-JUL-1999; 99US-0143907.
PR	27-AUG-1999; 99US-0151140.
PR	23-FEB-2000; 2000US-0184285.
PR	13-JUL-2000; 2000US-0610132.
XX	
XX	(UYLA-) UNIV LAUSANNE.

XX Thorens B, Ibberson M, Uldry M;
PI
XX
DR WP1: 2001-112615/12.
DR P-PSDB: AAB66933.
XX
PT Nucleic acids encoding GLUTx glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
PT e.g. Ischemia and diabetes -
XX
PS
PS Claim 3: Page 71-73; 124pp: English.
XX
CC The present invention relates to GLUTx proteins (AA655865-AA655871 and
CC AA66632-AA666941). The GLUTx proteins are related to the facultative
CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTx proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as Ischaemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is the coding sequence
CC for rat GLUTx1.
XX
XQ Sequence 2087 BP; 355 A; 673 C; 560 G; 479 T; 0 other;

Query Match	100.0%;	Score 282;	DB 22;	Length 2087;
Best Local Similarity	100.0%;	Pred. No. 2.8e-134;		
Matches 282; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

Qy	1	gagcctctcatcttcaaccctcgagctcgctgcctgcgcgtatgagcatcttgctcttc	60
Db	1116	gagcctctcatcttcaaccctcgagctcgctgcctgcgcgtatgagcatcttgctcttc	1175
Qy	61	atgcctggttttcagtaaggctcgaggaccatccctctgctctctatctatctagagatcttc	120
Db	1176	atgcctggttttcagtaaggctcgaggaccatccctctgctctctatctatctagagatcttc	1235
Qy	121	cctctgtaactaaagaagtttggtgtctaccggcgtctgtgtctcaaccaatggtttatgagcc	180
Db	1236	cctctgtaactaaagaagtttggtgtctaccggcgtctgtgtctcaaccaactgtttatgagcc	1235
Qy	181	tctctgtgtgaccaaagaagtttaacagcatcatatgagatcctcagaaccttaagcgctctc	240
Db	1296	tctctgtgtgaccaaagaagtttaacagcatcatatgagatcctcagaaccttaagcgctctc	1355
Qy	241	tgtgtcaaccgctgctctctgtatcaccctcagaagctctcttttcaag	282
Db	1356	tgtgtcaaccgctgctctctgtatcaccctcagaagctctcttttcaag	1397

RESULT	2
AAFS5867	
ID	AAFS5867 standard; cDNA; 2072 BP.
XX	
AC	AAFS5867;
XX	
DT	17-APR-2001 (first entry)
XX	
DE	Murine GLUTX1 coding sequence.
XX	
KW	Murine; GLUTX; gene therapy; vaccine; hexose transport modulator;
KW	hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
KW	hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX	
OS	Mus sp.
XX	
PN	WO200104145-A2.
XX	
PD	18-JAN-2001.
XX	
PF	14-JUL-2000; 2000MO-IB01042.
XX	
PR	14-JUL-1999; 990S-0143907.
PR	27-AUG-1999; 990S-0151140.

PR	23-FEB-2000;	2000US-0184285.	
PR	13-JUL-2000;	2000US-0616132.	
PA	(UYLA-) UNIV LAUSANNE.		
XX			
XX	Thorens B, Ibberson M, Uldry M;		
XX			
DR	WPI: 2001-112615/12.		
XX	P-PSDB: AAB66934.		
PT	Nucleic acids encoding GLUTX glucose transporter proteins, useful in		
PT	the prevention, diagnosis and treatment of hexose transport disorders,		
PT	e.g. ischemia and diabetes -		
XX			
PS	Claim 3; Page 73-74; 124pp; English.		
XX			
CC	The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and		
CC	AAB66932-AAB66941). The GLUTX proteins are related to the facultative		
CC	glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport and		
CC	function. The GLUTX proteins may be used in the diagnosis, prevention and		
CC	treatment of hexose transport disorders such as ischemia, diabetes,		
CC	hyperglycemia, hypoglycaemia, a glucose metabolism disorder and/or a		
CC	neurodegenerative disease. The present sequence is the coding sequence		
CC	for murine GLUTX1.		
SO			
XX	Sequence 2072 BP: 351 A; 673 C; 584 G; 464 T; 0 other;		
	Query Match	17.7%; Score 50; DB 22; Length 2072;	
	Best Local Similarity	100.0%; Pred. NO. 5.6e-16;	
	Matches	50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	79	gacttggggaccatccctgctccctcatgtcagaagatcttcctctgcga	128
Db	1182	gacttgaggaccatccctgctccctcatgtcagaagatcttcctctgcga	1231
	RESULT		
	3		
	AAAA4914/C		
	ID	AAAA4914 standard; CDNA: 579 BP.	
XX			
AC	AAAA4914;		
XX			
DT	21-AUG-2000 (first entry)		
XX			
DE	Human secreted expressed sequence tag SEQ ID NO:1489.		
XX			
KW	Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;		
KW	expressed sequence tag; EST; probe; chemotactic; proliferative;		
KW	immunomodulatory; haematopoietic; chemokine; analgesic; haemostatic;		
KW	thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;		
KW	antiviral; antidiabetic; antisthmatic; vulnery; antiparkinsonian;		
KW	antitumor; osteopathic; neuroprotective; nocotropic; antipsoriatic;		
KW	cerebroprotective; anticonvulsant; antidepressant; gene therapy;		
KW	vaccine; autoimmune disorder; multiple sclerosis; allergic condition;		
KW	insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;		
KW	lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;		
KW	central nervous system disorder; Alzheimer's disease; stroke;		
KW	Parkinson's disease; Huntington's disease; coagulation disorder;		
KW	haemophilia; thrombosis; inflammatory disorder; Crohn's disease;		
KW	tumour; infection; depression; psoriasis; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200021991-A1.		
XX			
PD	20-APR-2000.		
XX			
PF	15-OCT-1999; 99WO-US24206.		
XX			
PR	15-OCT-1998; 98US-0104436.		
XX			
PA	(GENY) GENETICS INST INC.		

XX	XX	Jacobs K, McCoy JM, LaValle ER, Collins-Racie LA, Evans C:
PI	PI	Merberg D, Treacy M, Bowman MR:
XX	XX	WP1: 2000-317938/27.
XX	XX	
PT	PT	Isolated polynucleotides, and encoded proteins, comprising secreted
PT	PT	expressed sequence tags (sESTs), useful for treating various disorders
XX	XX	such as autoimmune, infectious, and central nervous system disorders -
PS	PS	Claim 1: Page 586: 803pp: English.
XX	XX	
CC	CC	AAA43426 to AAA5925 represent specifically claimed secreted expressed
CC	CC	sequence tags (sESTs), isolated from human, mouse, chicken and rat
CC	CC	tissue sources. The sESTs can have a range of activities depending on
CC	CC	the tissues they were isolated from. The activities include:
CC	CC	chemotactic; proliferative; immunomodulatory; haematopoietic;
CC	CC	chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC	CC	cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
CC	CC	antihistaminic; vulnery; antitumor; osteoprotective; neuroprotective;
CC	CC	neurotropic; antiparkinsonian; antipsoriatic; cerebroprotective;
CC	CC	anticonvulsant; and antidepressant. The sESTs can be used for gene
CC	CC	therapy and in vaccines. The sESTs are useful as probes for the
CC	CC	identification and isolation of full-length cDNAs and genomic DNA
CC	CC	molecules which correspond to the sESTs. Proteins encoded by the sESTs
CC	CC	are useful in assays for determining biological activity and raising
CC	CC	antibodies. They may be useful for treatment of autoimmune disorders
CC	CC	(multiple sclerosis, insulin dependent diabetes), allergic conditions
CC	CC	(asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC	CC	osteoporosis, osteoarthritis, central nervous system disorders
CC	CC	(Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC	CC	disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC	CC	disease), tumours, bacterial, fungal or viral infections, depression and
CC	CC	psoriasis. AAA45926 to AAA45931 represent linker variants which are given
CC	CC	in the exemplification of the present invention.
XX	XX	
SO	SO	Sequence 579 BP: 109 A; 184 C; 180 G; 106 T; 0 other:
		Query Match 14.5%; Score 41; DB 21; Length 579;
		Best Local Similarity 100.0%; Pred. No. 2,1e-11;
		Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY	QY	88 cccatcccttgctcctcatgtcagagatcttcctctgcga 128
Db	Db	330 CCCATCCCCCTGGCTCCTCATGTCACAGATCTTCCCTGCA 290
RESULT 4		
ID	ID	AAAD09552
XX	XX	AAAD09552 standard; cDNA: 2080 BP.
AC	AC	AAAD09552:
XX	XX	
DT	DT	10-SEP-2001 (first entry)
XX	XX	
DE	DE	Human transporter and ion channel-1 (TRICH-1) cDNA.
XX	XX	
KM	KM	Human; transporter and ion channel-1, TRICH-1; vaccine: cystic fibrosis;
KM	KM	gene therapy: amyotrophic lateral sclerosis; amnesia; muscular dystrophy;
KM	KM	hypertension; angina; neurological disorder; asthma; bipolar disorder;
KM	KM	dementia; depression; Alzheimer's disease; epilepsy; mood: arhythmia;
KM	KM	Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke;
KM	KM	Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy;
KM	KM	demyelinating disease; mental disorder; Schizophrenia; polymyositis;
KM	KM	muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease;
KM	KM	dermatomyositis; diabetes mellitus; immunological disorder; psoriasis;
KM	KM	rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus;
KM	KM	sickle cell anaemia; Wilson's disease; infertility; Cushing's disease;
KM	KM	scleroderma; pulmonary artery stenosis; neurotropic; Addison's disease;
KM	KM	malabsorption syndrome; hypercholesterolaemia; cancer; ss.
XX	XX	
OS	OS	Homo sapiens

KW contraceptive; antiinfectility; antiinflammatory; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 61..849
FT /*tag= a
FT /product= "Human protein having hydrophobic domain"
FT /note= "CDS is specifically is claimed in claim 3"
XX
PN WO200149728-A2.
PD 12-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-JP09359.
XX
PR 06-JAN-2000; 2000JP-0000585.
PR 06-JAN-2000; 2000JP-0000588.
PR 11-JAN-2000; 2000JP-0002299.
PR 03-FEB-2000; 2000JP-0026862.
PR 03-MAR-2000; 2000JP-0058367.
XX
PA (PROT-) PROTEGENE INC.
PA (SAGA) SAGAMI CHEM RES. CENT.
PI Kato S, Kimura T;
XX
DR WPI: 2001-418355/44.
DR P-PSDB; AAE06579.
XX
PT Human proteins with hydrophobic domains and the nucleic acids encoding
PT them, useful for preventing diagnosing and treating e.g. cancer,
PT Alzheimer's and inflammation -
XX
PS Claim 4; Page 287-289; 563pp; English.
XX
CC The present sequence is human protein with hydrophobic domain encoding
CC cDNA clone HP10784. The polynucleotide and polypeptide of the invention
CC may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The polynucleotides
CC may be used to produce the polypeptide, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the protein. The
CC polynucleotides and its complementary sequences may also be used as DNA
CC probes in diagnostic assays and also used in gene therapy. The
CC polypeptides may also be used as antigens in the production of antibodies
CC and in assays to identify modulators of polypeptide expression and
CC activity. The polypeptides and nucleic acids may be used as nutritional
CC supplements, to modulate cytokine and cell proliferation activity, to
CC modulate immune stimulation or suppression (e.g. for the treatment of
CC microbial infections and autoimmune disorders such as multiple sclerosis,
CC rheumatoid arthritis and insulin-dependent diabetes), to modulate
CC haematopoiesis, to modulate tissue growth activity (e.g. for the
CC treatment of Parkinson's disease, Huntington's disease and Alzheimer's
CC disease), to modulate activin and inhibin activity (e.g. for controlling
CC fertility), to modulate chemotactic and chemokinetic activity, to
CC modulate haemostatic and thrombolytic activity, to modulate receptor
CC ligand activity, to modulate inflammation and to inhibit tumour growth.
XX
SQ Sequence 1461 BP; 195 A; 501 C; 456 G; 309 T; 0 other;

Query Match 8.2%; Score 23; DB 22; Length 1461;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 ggcagcatgtgccttcacgc 65
|||||
Db 881 ggcagcatgtgccttcacgc 903

RESULT 7
AA114336/C
ID AA114336 standard; DNA: 423 BP.

XX
AC AA114336;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #4269 for gene expression analysis in human cervical cell sample.
XX
KW Probe: human: microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488901/53.
XX
PT human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 25; SEQ ID NO 4269; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 423 BP; 104 A; 128 C; 109 G; 82 T; 0 other;

Query Match 7.1%; Score 20; DB 22; Length 423;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 tcacctggggctgcctcgc 34
|||||
Db 176 TCACCTGGGCTGCCTCGGC 157

RESULT 8
AA135713/C
ID AA135713 standard; DNA: 423 BP.
XX
AC AA135713;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #4399 used to measure gene expression in human placenta sample.
XX
KW Probe: microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX

OS Homo sapiens.
 XX
 PN WO200157272-A2.
 XX
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00663.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488897/53.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 PS Claim 25; SEQ ID No 4399; 654pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 XX
 SQ Sequence 423 BP; 104 A; 128 C; 109 G; 82 T; 0 other;

Query Match 7.1%; Score 20; DB 22; Length 423;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 15 tcactgggctgctgctgc 34
 ||||||||||||||||||||
 Db 176 TCACCTGGGGCTGGCTGGC 157
 RESULT 9
 AA104170/C
 ID AA104170 standard; DNA: 423 BP.
 XX
 AC AA104170;
 XX
 DT 09-OCT-2001 (first entry)
 XX
 DE Probe #4161 used to measure gene expression in human breast sample.
 XX
 KW Probe; human; breast disease; breast cancer; development disorder; ss;
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
 XX
 OS Homo sapiens.
 XX
 PM WO200157270-A2.
 PD
 XX 09-AUG-2001.
 PF
 XX 29-JAN-2001; 2001WO-US00661.
 PR
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-476286/51.
 DR
 XX Novel single exon nucleic acid probe used to measuring gene expression
 PT in a human breast -
 XX
 PS Claim 25; SEQ ID No 4161; 322pp; English.
 XX
 CC The present invention relates to novel single exon nucleic acid probes.
 CC The present sequence is one such probe. The probes are useful for
 CC measuring human gene expression in a human breast sample, where the probe
 CC hybridises at high stringency to a nucleic acid expressed in the human
 CC breast. The probes are useful for predicting, diagnosing, grading,
 CC staging, monitoring and prognosing diseases of the human breast,
 CC particularly those diseases with polygenic aetiology. The diseases
 CC include: breast cancer, disorders of development, inflammatory diseases
 CC of the breast, fibrocystic changes, proliferative breast disease and
 CC non-carcinoma tumours.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 423 BP; 104 A; 128 C; 109 G; 82 T; 0 other;

Query Match 7.1%; Score 20; DB 22; Length 423;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 15 tcactgggctgctgctgc 34
 ||||||||||||||||||||
 Db 176 TCACCTGGGGCTGGCTGGC 157
 RESULT 10
 AA123539/C
 ID AA123539 standard; DNA: 465 BP.
 XX
 AC AA123539;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Probe #13472 for gene expression analysis in human cervical cell sample.
 XX
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PM WO200157278-A2.
 PD
 XX 09-AUG-2001.
 PF
 XX 30-JAN-2001; 2001WO-US00670.
 PR
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488901/53.

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 25; SEQ ID No 13472; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human Hela cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SO Sequence 465 BP; 113 A; 137 C; 136 G; 79 T; 0 other;

Query Match 7.18; Score 20; DB 22; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 tcacctgggctggccctggc 34
|||||
DB 62 TCACCTGGGCTGGCTGGC 43

RESULT 11
AA148856/c
ID AA148856 standard; DNA: 465 BP.
XX
XX AA148856;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #17542 used to measure gene expression in human placenta sample.
XX
XX Probe; microarray: human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX
XX WO200157272-A2.
XX
XX
XX 09-AUG-2001.
XX
XX
XX 30-JAN-2001; 2001WO-US00663.
XX
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000GB-0024263.
XX
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488897/53.
XX
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX
XX
XX Claim 25; SEQ ID No 17542; 654pp; English.
XX
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
XX The present sequence is one such probe. The probes are useful for
XX producing a microarray for predicting, measuring and displaying gene
XX expression in samples derived from human placenta. The probes are useful
XX for antenatal diagnosis of human genetic disorders.

XX
SO Sequence 465 BP; 113 A; 137 C; 136 G; 79 T; 0 other;

Query Match 7.18; Score 20; DB 22; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 tcacctgggctggccctggc 34
|||||
DB 62 TCACCTGGGCTGGCTGGC 43

RESULT 12
AA109165/c
ID AA109165 standard; DNA: 465 BP.
XX
XX AA109165;
XX
DT 09-OCT-2001 (first entry)
XX
DE Probe #9156 used to measure gene expression in human breast sample.
XX
XX
XX Probe; human; breast disease; breast cancer; development disorder; ss;
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
XX Homo sapiens.
XX
XX
XX WO200157270-A2.
XX
XX
XX 09-AUG-2001.
XX
XX
XX 29-JAN-2001; 2001WO-US00661.
XX
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000GB-0024263.
XX
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-476286/51.
XX
XX
XX Novel single exon nucleic acid probe used to measuring gene expression
XX in a human breast -
XX
XX
XX Claim 25; SEQ ID No 9156; 322pp; English.
XX
XX
XX The present invention relates to novel single exon nucleic acid probes.
XX The present sequence is one such probe. The probes are useful for
XX measuring human gene expression in a human breast sample, where the probe
XX hybridises at high stringency to a nucleic acid expressed in the human
XX breast. The probes are useful for predicting, diagnosing, grading,
XX staging, monitoring and prognosing diseases of the human breast,
XX particularly those diseases with polygenic aetiology. The diseases
XX include: breast cancer, disorders of development, inflammatory diseases
XX of the breast, fibrocystic changes, proliferative breast disease and
XX non-carcinoma tumours.
XX
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
XX Sequence 465 BP; 113 A; 137 C; 136 G; 79 T; 0 other;

Query Match 7.18; Score 20; DB 22; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 tcacctgagagctgcctgac 34
 |||
 Db 62 TCACCTGGGGCTGCCTGC 43

RESULT 13
 AAX30839/c
 ID AAX30839 standard; DNA; 954 BP.
 XX
 AC AAX30839;
 XX

DT 20-MAY-1999 (first entry)

XX Streptococcus pneumoniae genomic DNA sequence SEQ ID NO:116.

XX Streptococcus pneumoniae strain 0100993; vaccine; immune response;
 KW Streptococcal infection; pneumococcal; ss.

XX Streptococcus pneumoniae.

XX OS W09737026-A1.

XX PD 09-OCT-1997.

XX PF 01-APR-1997; 97WO-US05306.

XX PR 22-AUG-1996; 96US-0025788.

XX PR 02-APR-1996; 96US-0014690.

XX PA (SMIK) SMITHKLINE BEECHAM CORP.

XX PA (SMIK) SMITHKLINE BEECHAM PLC.

XX PI Black MT, Hodgson JE, Knowles DJC, Nicholas RO;

XX PI Stodola RK;

XX DR WPI: 1997-503111/46.

XX DR P-PSDB; AAY11256.

XX PT Nucleic acids encoding pneumococcal polypeptide(s) - useful in

XX PT vaccines, drug screening, etc

XX PS Claim 5; Page 128; 354pp; English.

XX CC AAX30724 to AAX30946 represent genomic DNA sequences isolated from

XX CC Streptococcus pneumoniae strain 0100993. These genomic DNA sequences

XX CC encode the novel proteins given in AAY1114 to AAY11367. The proteins,

XX CC isolated from Streptococcus pneumoniae, can be used in vaccines against

XX CC streptococcal infections and in assays for identifying compounds that

XX CC inhibit or activate the activity of the proteins. The antagonists can

XX CC be used to treat an individual having need to inhibit a bacterial

XX CC protein. Vectors expressing the proteins can be used to induce a

XX CC protective immune response in mammals.

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX Streptococcus pneumoniae coding region.
 DE
 XX coding region; ORF; open reading frame; antibacterial;
 KW infection; prevention; meningitis; ss.
 KW
 XX Streptococcus pneumoniae.
 OS

XX Key location/Qualifiers

XX CDS 687..1055

XX FT /*tag= a

XX FT /product= unknown

XX FT 979..1932

XX FT /*tag= b

XX FT /product= unknown

XX W09819689-A1.

XX 14-MAY-1998.

XX 27-OCT-1997; 97WO-US19226.

XX 01-NOV-1996; 96US-0029930.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX (SMIK) SMITHKLINE BEECHAM PLC.

XX Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;

XX Reid RH, Zarfos PN;

XX WPI: 1998-286586/25.

XX P-PSDB; AAW60996; AAW60997.

XX New isolated nucleic acids from Streptococcus pneumoniae - useful,

XX e.g. for identifying anti-bacterial(s) for treatment and prevention

XX of meningitis

XX Claim 1; Page 91-92; 130pp; English.

XX The sequence is that of a coding region isolated from

XX S. pneumoniae. Its encoded protein, or agonists of it,

XX may be useful as an antibacterial for treatment or

XX prevention of infection, specifically caused by S. pneumoniae

XX (particularly meningitis) but possibly also Helicobacter

XX pylori (ulcers and gastric cancer). It may be of particular

XX use before insertion of an in-dwelling device or any other

XX invasive procedure. The protein, or nucleic acid encoding

XX it, can also be used in vaccines to induce a cellular

XX and/or humoral immune response, or to screen for other

XX antibacterials. The DNA may also contain flanking sequences

XX that are potential sources of control elements for bacterial

XX gene expression. Detecting a sequence encoding the protein

XX can be used diagnostically, e.g. to detect a mutation for

XX serotyping or classifying infectious agents.

XX Sequence 2051 BP; 569 A; 358 C; 458 G; 666 T; 0 other;

XX Query Match 6.7%; Score 19; DB 19; Length 2051;

XX Best Local Similarity 100.0%; Pred. No. 3.6;

XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 255 ctctgtatcctcagcgtc 273
 |||
 Db 1609 CTCTGTATCCTCAGCGTC 1591

RESULT 14
 AAV37386/c

ID AAV37386 standard; DNA; 2051 BP.

XX AAV37386;

XX 13-OCT-1998 (first entry)

XX

XX

XX

DT 10-SEP-2001 (first entry)
 DE Human heart cell specific cDNA sequence SEQ ID NO:215.
 XX
 XX
 KW Human: tissue specific: diagnosis: brain; heart; skeletal muscle;
 KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
 KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
 KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
 XX
 OS Homo sapiens.
 XX
 PN WO200132927-A2.
 XX
 PD 10-MAY-2001.
 XX
 XX
 PF 02-NOV-2000: 2000WO-US30396.
 XX
 PR 04-NOV-1999: 99US-0163508.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Sornasse T, Selhamer JJ, Watson GA;
 DR WPI; 2001-291057/30.
 XX
 PT New cell and tissue specific polynucleotides useful for diagnosis,
 PT prognosis or monitoring of treatments for disorders where the gene is
 PT associated with a cancer, immunopathology or neuropathology -
 XX
 PS Claim 1; Page 145-146; 327pp; English.
 XX
 CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
 CC sequences (I). (I) can have cytostatic, immunomodulatory and
 CC neuroprotective activities, and can be used in gene therapy. (I) and
 CC proteins (II) encoded by then are used in high throughput screening
 CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
 CC mimetics, peptides, proteins, agonists, antagonists, antibodies or
 CC their fragments, immunoglobulins, inhibitors, drug compounds and
 CC pharmaceutical agents. Expression of (I) in a sample indicates the
 CC differentiation of embryonic stem cells into a tissue selected from
 CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
 CC tissues. (I) and (II) are used to produce an expression profile that
 CC defines a metabolic or developmental process, treatment, condition,
 CC disease or disorder. The gene profile can be used for diagnosis,
 CC prognosis or monitoring of treatments and for investigating a
 CC predisposition to a disorder where the gene is associated with a
 CC cancer, immunopathology or neuropathology.
 XX
 SQ Sequence 4214 BP: 930 A; 1213 C; 1343 G; 727 T; 1 other:

Query Match 6.7%; Score 19; DB 22; Length 4214;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 19 ctgggagctgacctgactg 37
 ||||||||||||||||
 Db 2147 CTGGGCTGGCTGGCTGG 2129

Search completed: February 13, 2002, 20:10:52
 Job time: 13504 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 20:05:31 : Search time 130.06 Seconds
(without alignments)
491.056 Million cell updates/sec

Title: US-09-516-493-11
Perfect score: 282
Sequence: 1 gagctgtgtatgtaccct.....tccacagctccttcacag 282

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 351203 seqs, 113238999 residues

Word size : 0

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
4: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	6.4	2831	2 US-08-808-982-3	Sequence 3, Appli
2	18	6.4	2831	4 US-09-306-902A-3	Sequence 3, Appli
3	16	5.7	569	1 US-08-036-555B-44	Sequence 44, Appli
4	16	5.7	569	1 US-08-469-569-44	Sequence 44, Appli
5	16	5.7	569	1 US-08-249-322A-44	Sequence 44, Appli
6	16	5.7	569	1 US-08-469-526A-44	Sequence 44, Appli
7	16	5.7	569	2 US-08-734-591A-44	Sequence 44, Appli
8	16	5.7	569	2 US-08-469-660-181	Sequence 181, App
9	16	5.7	569	3 US-08-341-018-43	Sequence 43, Appli
10	16	5.7	569	3 US-08-470-335-44	Sequence 44, Appli
11	16	5.7	569	4 US-08-735-021-44	Sequence 44, Appli
12	16	5.7	569	4 US-08-734-664A-44	Sequence 44, Appli
13	16	5.7	569	4 US-08-470-339-44	Sequence 44, Appli
14	16	5.7	569	5 PCT-US94-05083C-177	Sequence 177, App
15	16	5.7	569	5 PCT-US95-06846A-44	Sequence 177, App
16	16	5.7	631	4 US-08-943-731-187	Sequence 187, App
17	16	5.7	959	1 US-08-634-331-1	Sequence 1, Appli
18	16	5.7	959	6 5508165-1	Patent No. 5508165
19	16	5.7	1351	4 US-08-697-954-3	Sequence 3, Appli
20	16	5.7	1403	3 US-09-013-881-10	Sequence 10, Appli
21	16	5.7	1637	4 US-09-160-494-5	Sequence 5, Appli
22	16	5.7	1899	1 US-08-634-331-2	Sequence 2, Appli
23	16	5.7	2010	1 US-07-847-743B-12	Sequence 12, Appli
24	16	5.7	2010	1 US-08-456-201-12	Sequence 12, Appli
25	16	5.7	2010	2 US-08-330-161-10	Sequence 10, Appli
26	16	5.7	2010	2 US-08-456-241-12	Sequence 12, Appli
27	16	5.7	2010	2 US-08-440-401-10	Sequence 10, Appli

C 28	16	5.7	2010	2 US-08-419-878B-10	Sequence 10, Appli
C 29	16	5.7	2010	5 PCT-US92-04295A-12	Sequence 12, Appli
C 30	16	5.7	2164	1 US-08-096-277-6	Sequence 6, Appli
C 31	16	5.7	2164	1 US-08-550-815-6	Sequence 6, Appli
C 32	16	5.7	2164	3 US-08-703-089-6	Sequence 6, Appli
C 33	16	5.7	2199	1 US-07-847-743B-7	Sequence 7, Appli
C 34	16	5.7	2199	1 US-08-096-277-7	Sequence 7, Appli
C 35	16	5.7	2199	1 US-08-456-201-7	Sequence 7, Appli
C 36	16	5.7	2199	2 US-08-456-241-7	Sequence 7, Appli
C 37	16	5.7	2199	2 US-08-550-815-7	Sequence 7, Appli
C 38	16	5.7	2199	3 US-08-703-089-7	Sequence 7, Appli
C 39	16	5.7	2199	5 PCT-US92-04295A-7	Sequence 7, Appli
C 40	16	5.7	2351	1 US-08-168-091A-1	Sequence 1, Appli
C 41	16	5.7	2431	1 US-07-847-743B-25	Sequence 25, Appli
C 42	16	5.7	2431	1 US-08-456-201-25	Sequence 25, Appli
C 43	16	5.7	2431	1 US-08-456-241-25	Sequence 25, Appli
C 44	16	5.7	2431	5 PCT-US92-04295A-25	Sequence 25, Appli
C 45	16	5.7	2490	1 US-07-847-743B-23	Sequence 23, Appli
C 46	16	5.7	2490	1 US-08-456-201-23	Sequence 23, Appli
C 47	16	5.7	2490	1 US-08-456-241-23	Sequence 23, Appli
C 48	16	5.7	2490	5 PCT-US92-04295A-23	Sequence 23, Appli
C 49	16	5.7	2545	1 US-07-869-933-22	Sequence 22, Appli
C 50	16	5.7	2545	4 US-09-103-663-22	Sequence 22, Appli
C 51	16	5.7	3212	4 US-08-697-954-1	Sequence 1, Appli
C 52	16	5.7	4332	3 US-08-728-603-14	Sequence 14, Appli
C 53	16	5.7	5143	1 US-08-574-043A-7	Sequence 7, Appli
C 54	16	5.7	5143	2 US-08-795-015-7	Sequence 7, Appli
C 55	16	5.7	20084	4 US-08-943-731-5	Sequence 5, Appli
C 56	16	5.7	26700	1 US-08-472-217-1	Sequence 1, Appli
C 57	16	5.7	26700	2 US-08-488-199-5	Sequence 5, Appli
C 58	16	5.7	26700	3 US-08-760-534A-1	Sequence 1, Appli
C 59	16	5.7	32207	2 US-08-770-379-9	Sequence 20, Appli
C 60	16	5.7	32207	4 US-08-757-669A-20	Sequence 20, Appli
C 61	15	5.3	27	2 US-08-696-770-4	Sequence 4, Appli
C 62	15	5.3	27	2 US-09-015-557-4	Sequence 4, Appli
C 63	15	5.3	281	1 US-08-262-424-3	Sequence 3, Appli
C 64	15	5.3	504	2 US-08-747-887-22	Sequence 22, Appli
C 65	15	5.3	514	3 US-08-875-573-5	Sequence 5, Appli
C 66	15	5.3	646	4 US-09-328-111-658	Sequence 658, App
C 67	15	5.3	672	1 US-08-485-455D-16	Sequence 16, Appli
C 68	15	5.3	672	2 US-08-485-455D-16	Sequence 16, Appli
C 69	15	5.3	672	2 US-08-484-211C-16	Sequence 16, Appli
C 70	15	5.3	672	3 US-08-906-616-16	Sequence 16, Appli
C 71	15	5.3	672	3 US-08-906-616-16	Sequence 16, Appli
C 72	15	5.3	672	3 US-08-817-795-16	Sequence 16, Appli
C 73	15	5.3	672	3 US-08-485-443B-16	Sequence 16, Appli
C 74	15	5.3	672	3 US-08-639-075A-16	Sequence 16, Appli
C 75	15	5.3	672	4 US-09-012-631-16	Sequence 16, Appli
C 76	15	5.3	672	4 US-09-012-692-16	Sequence 16, Appli
C 77	15	5.3	672	4 US-08-906-613-16	Sequence 16, Appli
C 78	15	5.3	672	5 PCT-US95-14442A-16	Sequence 16, Appli
C 79	15	5.3	854	1 US-08-592-126-76	Sequence 76, Appli
C 80	15	5.3	1124	2 US-08-742-440A-2	Sequence 2, Appli
C 81	15	5.3	1131	6 5168049-2	Patent No. 5168049
C 82	15	5.3	1210	4 US-08-811-481-30	Sequence 30, Appli
C 83	15	5.3	1235	1 US-08-696-770-1	Sequence 1, Appli
C 84	15	5.3	1235	1 US-09-015-557-1	Sequence 1, Appli
C 85	15	5.3	1347	4 US-09-050-159-124	Sequence 124, App
C 86	15	5.3	1349	1 US-07-951-715A-20	Sequence 20, Appli
C 87	15	5.3	1349	2 US-08-459-448B-20	Sequence 20, Appli
C 88	15	5.3	1349	3 US-08-459-595A-20	Sequence 20, Appli
C 89	15	5.3	1349	3 US-08-459-504A-20	Sequence 20, Appli
C 90	15	5.3	1349	3 US-08-459-444-20	Sequence 20, Appli
C 91	15	5.3	1350	4 US-09-232-190-62	Sequence 62, Appli
C 92	15	5.3	1350	4 US-09-232-197-62	Sequence 62, Appli
C 93	15	5.3	1435	2 US-08-239-276-9	Sequence 9, Appli
C 94	15	5.3	1435	2 US-08-468-577B-9	Sequence 9, Appli
C 95	15	5.3	1413	3 US-08-468-577B-9	Sequence 9, Appli
C 96	15	5.3	1419	4 US-09-276-531-1	Sequence 1, Appli
C 97	15	5.3	1496	4 US-09-384-212-1	Sequence 19, Appli
C 98	15	5.3	1607	3 US-08-875-573-19	Sequence 19, Appli
C 99	15	5.3	1650	2 US-08-436-771-10	Sequence 10, Appli
C 100	15	5.3	1650	2 US-08-434-998-10	Sequence 10, Appli

c 101	15	5.3	1650	2	US-08-487-797-10	Sequence 10, Appl	c 174	14	5.0	569	3	US-08-470-335-146	Sequence 16, App
c 102	15	5.3	1650	5	PCT-US95-02058-10	Sequence 10, Appl	c 175	14	5.0	569	4	US-08-735-021-146	Sequence 16, App
c 103	15	5.3	1695	4	US-09-233-878-1	Sequence 1, Appl	c 176	14	5.0	569	4	US-08-734-664A-146	Sequence 14, App
c 104	15	5.3	1983	4	US-09-057-860A-8	Sequence 8, Appl	c 177	14	5.0	569	4	US-08-470-339-146	Sequence 14, App
c 105	15	5.3	2005	1	US-08-208-108-1	Sequence 1, Appl	c 178	14	5.0	569	5	PCT-US94-05083C-142	Sequence 14, App
c 106	15	5.3	2076	5	PCT-US91-08442-1	Sequence 1, Appl	c 179	14	5.0	569	5	PCT-US95-06846A-146	Sequence 16, App
c 107	15	5.3	2201	1	US-08-453-472-2	Sequence 2, Appl	c 180	14	5.0	588	2	US-08-924-695A-3	Sequence 3, Appl
c 108	15	5.3	2201	1	US-08-038-948-6	Sequence 6, Appl	c 181	14	5.0	600	1	US-08-233-146-3	Sequence 3, Appl
c 109	15	5.3	2201	1	US-08-453-952-2	Sequence 2, Appl	c 182	14	5.0	600	1	US-08-463-470-3	Sequence 3, Appl
c 110	15	5.3	2201	2	US-08-862-903-2	Sequence 2, Appl	c 183	14	5.0	655	4	US-09-385-982-140	Sequence 340, App
c 111	15	5.3	2345	3	US-08-955-937A-1	Sequence 1, Appl	c 184	14	5.0	737	1	US-07-872-678A-4	Sequence 4, Appl
c 112	15	5.3	2345	4	US-09-300-985-1	Sequence 1, Appl	c 185	14	5.0	789	2	US-08-602-359A-28	Sequence 28, Appl
c 113	15	5.3	2353	4	PCT-US92-06840-1	Sequence 1, Appl	c 186	14	5.0	827	2	US-09-333-599-5	Sequence 5, Appl
c 114	15	5.3	2367	3	US-08-276-968A-17	Sequence 17, Appl	c 187	14	5.0	836	1	US-08-554-612C-26	Sequence 26, Appl
c 115	15	5.3	2624	1	US-08-016-863-15	Sequence 15, Appl	c 188	14	5.0	870	4	US-09-333-599-1	Sequence 1, Appl
c 116	15	5.3	2624	4	US-08-276-968A-15	Sequence 15, Appl	c 189	14	5.0	892	2	US-08-460-529B-1	Sequence 1, Appl
c 117	15	5.3	3023	4	US-09-593-589-10	Sequence 10, Appl	c 190	14	5.0	906	3	US-08-831-132-1	Sequence 1, Appl
c 118	15	5.3	3113	2	US-08-993-228-20	Sequence 20, Appl	c 191	14	5.0	906	4	US-09-416-150-1	Sequence 1, Appl
c 119	15	5.3	3311	2	US-08-239-276-10	Sequence 10, Appl	c 192	14	5.0	911	1	US-08-745-603-1	Sequence 1, Appl
c 120	15	5.3	3311	2	US-08-468-579B-10	Sequence 10, Appl	c 193	14	5.0	932	1	US-08-554-612C-20	Sequence 20, Appl
c 121	15	5.3	3311	3	US-08-468-577B-10	Sequence 10, Appl	c 194	14	5.0	972	4	US-09-479-524-1	Sequence 1, Appl
c 122	15	5.3	3503	1	US-07-631-717A-1	Sequence 1, Appl	c 195	14	5.0	972	4	US-09-479-524-7	Sequence 7, Appl
c 123	15	5.3	3503	1	US-08-166-717D-1	Sequence 1, Appl	c 196	14	5.0	1007	1	US-08-554-612C-17	Sequence 17, Appl
c 124	15	5.3	4080	2	US-08-710-249-3	Sequence 3, Appl	c 197	14	5.0	1007	1	US-08-554-612C-18	Sequence 18, Appl
c 125	15	5.3	4080	2	US-09-220-157A-3	Sequence 3, Appl	c 198	14	5.0	1014	2	US-08-359-850-1	Sequence 1, Appl
c 126	15	5.3	4162	2	US-08-459-448A-26	Sequence 26, Appl	c 199	14	5.0	1062	1	US-07-946-497-3	Sequence 3, Appl
c 127	15	5.3	4162	3	US-08-459-595A-26	Sequence 26, Appl	c 200	14	5.0	1062	1	US-08-483-332-3	Sequence 3, Appl
c 128	15	5.3	4162	3	US-08-459-504B-26	Sequence 26, Appl	c 201	14	5.0	1062	2	US-08-478-882-3	Sequence 3, Appl
c 129	15	5.3	4162	3	US-08-459-444-26	Sequence 26, Appl	c 202	14	5.0	1086	1	US-08-361-708-2	Sequence 2, Appl
c 130	15	5.3	4165	1	US-07-951-715A-26	Sequence 26, Appl	c 203	14	5.0	1086	1	US-08-536-277-2	Sequence 2, Appl
c 131	15	5.3	4300	1	US-08-041-538-1	Sequence 1, Appl	c 204	14	5.0	1243	1	US-08-314-309A-24	Sequence 24, Appl
c 132	15	5.3	4300	1	US-08-463-642-1	Sequence 1, Appl	c 205	14	5.0	1257	3	US-08-487-748A-11	Sequence 11, Appl
c 133	15	5.3	4300	1	US-08-455-602-1	Sequence 1, Appl	c 206	14	5.0	1257	3	US-08-480-070C-11	Sequence 11, Appl
c 134	15	5.3	4300	2	US-08-465-157-1	Sequence 1, Appl	c 207	14	5.0	1257	3	US-08-829-525-11	Sequence 11, Appl
c 135	15	5.3	4300	5	PCT-US91-09422-16	Sequence 1, Appl	c 208	14	5.0	1257	4	US-08-609-583A-11	Sequence 11, Appl
c 136	15	5.3	5236	5	US-08-965-762-1	Sequence 16, Appl	c 209	14	5.0	1257	4	US-08-937-399-11	Sequence 11, Appl
c 137	15	5.3	5596	4	US-08-716-351A-1	Sequence 1, Appl	c 210	14	5.0	1284	2	US-09-161-015-1	Sequence 1, Appl
c 138	15	5.3	8535	3	US-08-770-379-18	Sequence 18, Appl	c 211	14	5.0	1298	1	US-08-361-708-1	Sequence 1, Appl
c 139	15	5.3	35100	2	US-08-757-669A-18	Sequence 18, Appl	c 212	14	5.0	1298	1	US-08-536-277-1	Sequence 1, Appl
c 140	15	5.3	35100	4	US-09-036-987A-1	Sequence 1, Appl	c 213	14	5.0	1368	2	US-08-204-288-1	Sequence 1, Appl
c 141	15	5.3	80161	3	US-09-370-700-1	Sequence 1, Appl	c 214	14	5.0	1459	2	US-08-970-133-2	Sequence 2, Appl
c 142	15	5.3	80161	4	US-09-103-840A-2	Sequence 2, Appl	c 215	14	5.0	1503	2	US-08-715-325-1	Sequence 1, Appl
c 143	15	5.3	4403765	4	US-09-103-840A-1	Sequence 1, Appl	c 216	14	5.0	1508	3	US-08-554-612C-16	Sequence 16, Appl
c 144	15	5.3	4411529	4	US-09-103-840A-1	Sequence 1, Appl	c 217	14	5.0	1548	3	US-08-665-259-28	Sequence 28, Appl
c 145	14	5.0	52	3	US-08-875-573-7	Sequence 7, Appl	c 218	14	5.0	1548	3	US-08-762-500-28	Sequence 28, Appl
c 146	14	5.0	88	4	US-09-208-966-28	Sequence 28, Appl	c 219	14	5.0	1635	3	US-09-234-332-4	Sequence 4, Appl
c 147	14	5.0	102	3	US-08-564-225-7	Sequence 30, Appl	c 220	14	5.0	1673	1	US-07-791-936A-1	Sequence 1, Appl
c 148	14	5.0	165	1	US-08-324-243-26	Sequence 26, Appl	c 221	14	5.0	1673	1	US-08-383-781B-1	Sequence 1, Appl
c 149	14	5.0	165	1	US-08-532-390-26	Sequence 26, Appl	c 222	14	5.0	1681	1	US-07-791-936A-4	Sequence 4, Appl
c 150	14	5.0	165	3	US-08-717-294-26	Sequence 26, Appl	c 223	14	5.0	1681	1	US-08-383-781B-4	Sequence 4, Appl
c 151	14	5.0	165	5	PCT-US95-11511-26	Sequence 26, Appl	c 224	14	5.0	1683	1	US-07-791-936A-3	Sequence 3, Appl
c 152	14	5.0	211	4	US-09-269-617-8	Sequence 8, Appl	c 225	14	5.0	1763	1	US-08-383-781B-3	Sequence 3, Appl
c 153	14	5.0	249	1	US-08-330-108-6	Sequence 6, Appl	c 226	14	5.0	1764	1	US-08-036-555B-150	Sequence 150, App
c 154	14	5.0	249	5	PCT-US92-10087-6	Sequence 6, Appl	c 227	14	5.0	1764	1	US-08-469-569-150	Sequence 150, App
c 155	14	5.0	254	4	US-09-112-096-8	Sequence 8, Appl	c 228	14	5.0	1764	1	US-08-249-322A-150	Sequence 150, App
c 156	14	5.0	340	4	US-09-221-298-60	Sequence 60, Appl	c 229	14	5.0	1764	2	US-08-469-526A-150	Sequence 150, App
c 157	14	5.0	360	3	US-09-188-930-99	Sequence 99, Appl	c 230	14	5.0	1764	2	US-08-734-591A-150	Sequence 150, App
c 158	14	5.0	370	3	US-09-188-930-103	Sequence 103, App	c 231	14	5.0	1764	3	US-08-469-660-150	Sequence 150, App
c 159	14	5.0	375	3	US-08-870-370-7	Sequence 7, Appl	c 232	14	5.0	1764	3	US-08-341-018-57	Sequence 57, Appl
c 160	14	5.0	375	3	US-08-870-370-8	Sequence 8, Appl	c 233	14	5.0	1764	4	US-08-470-335-150	Sequence 150, App
c 161	14	5.0	375	3	US-08-870-370-9	Sequence 9, Appl	c 234	14	5.0	1764	4	US-08-734-664A-150	Sequence 150, App
c 162	14	5.0	386	4	US-08-905-223-171	Sequence 171, App	c 235	14	5.0	1764	4	US-08-470-339-150	Sequence 150, App
c 163	14	5.0	401	4	US-09-221-298-7	Sequence 7, Appl	c 236	14	5.0	1764	4	US-08-470-339-150	Sequence 150, App
c 164	14	5.0	401	4	PCT-US93-06251-92	Sequence 92, Appl	c 237	14	5.0	1764	5	PCT-US94-05083C-146	Sequence 146, App
c 165	14	5.0	423	4	US-09-221-298-2	Sequence 2, Appl	c 238	14	5.0	1764	5	PCT-US95-06846A-150	Sequence 150, App
c 166	14	5.0	423	4	US-08-036-555B-146	Sequence 146, App	c 239	14	5.0	1771	2	US-07-969-267B-1	Sequence 1, Appl
c 167	14	5.0	569	1	US-08-469-569-146	Sequence 146, App	c 240	14	5.0	1790	2	US-08-810-453-1	Sequence 1, Appl
c 168	14	5.0	569	1	US-08-249-322A-146	Sequence 146, App	c 241	14	5.0	1841	5	PCT-US95-00362-1	Sequence 1, Appl
c 169	14	5.0	569	1	US-08-469-526A-146	Sequence 146, App	c 242	14	5.0	1872	4	US-09-232-200-88	Sequence 88, Appl
c 170	14	5.0	569	2	US-08-734-591A-146	Sequence 146, App	c 243	14	5.0	1884	1	US-09-232-197-88	Sequence 88, Appl
c 171	14	5.0	569	2	US-08-469-660-146	Sequence 146, App	c 244	14	5.0	1884	1	US-07-704-288C-1	Sequence 1, Appl
c 172	14	5.0	569	3	US-08-469-660-146	Sequence 146, App	c 245	14	5.0	1884	1	US-08-379-259-1	Sequence 1, Appl
c 173	14	5.0	569	3	US-08-341-018-41	Sequence 41, Appl	c 246	14	5.0	1890	4	US-09-290-640-24	Sequence 24, Appl

247	14	5.0	1954	4	US-09-370-807-3	Sequence 3, Appl	C 320	14	5.0	4207	4	US-08-660-148-1	Sequence 1, Appl
248	14	5.0	2010	4	US-09-240-410-1	Sequence 1, Appl	C 321	14	5.0	4207	4	US-08-660-148-3	Sequence 4, Appl
249	14	5.0	2070	1	US-07-766-351-6	Sequence 6, Appl	C 322	14	5.0	4252	2	US-08-475-844-4	Sequence 3, Appl
250	14	5.0	2070	1	US-08-059-032-6	Sequence 6, Appl	C 323	14	5.0	4252	5	PCT-US95-08429-4	Sequence 4, Appl
251	14	5.0	2070	5	PCT-US91-07290-6	Sequence 6, Appl	C 324	14	5.0	4303	4	US-08-660-148-4	Sequence 4, Appl
252	14	5.0	2073	4	US-09-032-523-6	Sequence 6, Appl	C 325	14	5.0	4303	4	US-08-660-148-6	Sequence 6, Appl
253	14	5.0	2121	1	US-08-331-394-3	Sequence 3, Appl	C 326	14	5.0	4320	2	US-08-472-534-4	Sequence 4, Appl
254	14	5.0	2121	1	US-08-250-858-3	Sequence 3, Appl	C 327	14	5.0	5092	2	US-09-412-545-1	Sequence 1, Appl
255	14	5.0	2121	1	US-08-446-915-3	Sequence 3, Appl	C 328	14	5.0	5162	2	US-08-916-117-13	Sequence 1, Appl
256	14	5.0	2121	1	US-08-744-139-3	Sequence 3, Appl	C 329	14	5.0	5162	3	US-09-225-170-13	Sequence 13, Appl
257	14	5.0	2121	5	PCT-US95-06639-3	Sequence 3, Appl	C 330	14	5.0	5394	3	US-08-688-376-1	Sequence 1, Appl
258	14	5.0	2183	1	US-08-444-792-3	Sequence 3, Appl	C 331	14	5.0	5589	1	US-08-465-795-1	Sequence 1, Appl
259	14	5.0	2183	1	US-08-445-042-3	Sequence 3, Appl	C 332	14	5.0	5816	1	US-09-220-641-4	Sequence 4, Appl
260	14	5.0	2336	5	US-08-247-946A-1	Sequence 1, Appl	C 333	14	5.0	5843	1	US-08-554-612C-2	Sequence 1, Appl
261	14	5.0	2336	5	PCT-US95-06420-1	Sequence 1, Appl	C 334	14	5.0	5962	6	5386025-5	Patent No. 5386025
262	14	5.0	2387	4	US-08-288-408-4	Sequence 4, Appl	C 335	14	5.0	5975	1	US-08-404-354B-1	Sequence 1, Appl
263	14	5.0	2387	4	US-08-655-782-4	Sequence 4, Appl	C 336	14	5.0	5975	1	US-08-314-083B-1	Sequence 1, Appl
264	14	5.0	2415	3	US-09-306-922-1	Sequence 1, Appl	C 337	14	5.0	5975	1	US-08-435-675B-1	Sequence 1, Appl
265	14	5.0	2456	1	US-08-223-146-8	Sequence 8, Appl	C 338	14	5.0	5975	1	US-08-336-257A-3	Sequence 3, Appl
266	14	5.0	2456	1	US-08-463-470-8	Sequence 8, Appl	C 339	14	5.0	5975	3	US-08-884-599-1	Sequence 1, Appl
267	14	5.0	2464	2	US-08-941-262-2	Sequence 2, Appl	C 340	14	5.0	6727	3	US-08-629-643A-5	Sequence 5, Appl
268	14	5.0	2472	3	US-08-481-435-3	Sequence 3, Appl	C 341	14	5.0	6727	3	US-09-280-799-1	Sequence 1, Appl
269	14	5.0	2479	4	US-09-342-749-29	Sequence 29, Appl	C 342	14	5.0	6727	4	US-09-155-884-5	Sequence 5, Appl
270	14	5.0	2498	4	US-09-041-236-1	Sequence 1, Appl	C 343	14	5.0	8635	3	US-09-075-272-1	Sequence 1, Appl
271	14	5.0	2531	3	US-08-714-918-60	Sequence 60, Appl	C 344	14	5.0	8906	2	US-08-826-267-1	Sequence 1, Appl
272	14	5.0	2531	4	US-09-265-315-60	Sequence 60, Appl	C 345	14	5.0	9551	2	US-08-056-200-93	Sequence 93, Appl
273	14	5.0	2531	4	US-09-265-315-60	Sequence 60, Appl	C 346	14	5.0	9551	2	US-08-800-644-93	Sequence 93, Appl
274	14	5.0	2531	4	US-09-266-417-60	Sequence 60, Appl	C 347	14	5.0	11613	1	US-08-484-044-10	Sequence 10, Appl
275	14	5.0	2544	2	US-08-469-412A-6	Sequence 6, Appl	C 348	14	5.0	11832	2	US-08-416-603-1	Sequence 1, Appl
276	14	5.0	2544	4	US-09-021-715-6	Sequence 6, Appl	C 349	14	5.0	13146	2	US-08-724-354D-3	Sequence 3, Appl
277	14	5.0	2606	4	US-09-252-816A-7	Sequence 7, Appl	C 350	14	5.0	13146	3	US-09-270-984A-3	Sequence 3, Appl
278	14	5.0	2797	2	US-09-021-323-2	Sequence 2, Appl	C 351	14	5.0	13529	3	US-09-144-085-3	Sequence 3, Appl
279	14	5.0	2815	4	US-09-276-531-127	Sequence 127, App	C 352	14	5.0	35081	2	US-08-752-760A-1	Sequence 1, Appl
280	14	5.0	2882	1	US-08-424-567-1	Sequence 1, Appl	C 353	14	5.0	49272	1	US-08-614-770A-1	Sequence 1, Appl
281	14	5.0	2882	2	US-08-711-928-1	Sequence 1, Appl	C 354	14	5.0	56516	2	US-08-996-306-1	Sequence 1, Appl
282	14	5.0	2882	4	US-09-184-937-1	Sequence 1, Appl	C 355	14	5.0	56516	4	US-09-338-807-1	Sequence 1, Appl
283	14	5.0	2886	1	US-08-554-612C-51	Sequence 51, Appl	C 356	14	5.0	56520	4	US-09-338-907-179	Sequence 179, App
284	14	5.0	2909	1	US-08-554-612C-10	Sequence 10, Appl	C 357	14	5.0	152331	3	US-09-128-155-16	Sequence 16, Appl
285	14	5.0	2909	1	US-08-554-612C-11	Sequence 11, Appl	C 358	14	5.0	176373	3	US-09-128-155-17	Sequence 17, Appl
286	14	5.0	2946	1	US-08-473-122-1	Sequence 1, Appl	C 359	14	5.0	246240	2	US-08-724-394A-20	Sequence 20, Appl
287	14	5.0	2946	2	US-08-472-478-1	Sequence 1, Appl	C 360	14	5.0	246240	2	US-08-724-394A-21	Sequence 21, Appl
288	14	5.0	2946	2	US-08-463-081B-7	Sequence 7, Appl	C 361	14	5.0	246240	2	US-08-724-394A-22	Sequence 22, Appl
289	14	5.0	2946	2	US-08-461-379A-7	Sequence 7, Appl	C 362	13	4.6	18	4	US-09-282-147-12	Sequence 12, Appl
290	14	5.0	2946	2	US-08-462-390B-7	Sequence 7, Appl	C 363	13	4.6	20	2	US-08-904-901-79	Sequence 79, Appl
291	14	5.0	2946	3	US-08-463-074B-7	Sequence 7, Appl	C 364	13	4.6	20	3	US-09-249-730-79	Sequence 79, Appl
292	14	5.0	2946	3	US-08-465-585C-7	Sequence 7, Appl	C 365	13	4.6	20	4	US-09-061-764A-14	Sequence 14, Appl
293	14	5.0	2946	3	US-08-652-446-7	Sequence 7, Appl	C 366	13	4.6	22	1	US-08-853-203-17	Sequence 17, Appl
294	14	5.0	3065	1	US-08-156-866-1	Sequence 1, Appl	C 367	13	4.6	24	2	US-08-859-998-961	Sequence 961, App
295	14	5.0	3167	2	US-08-472-534-1	Sequence 1, Appl	C 368	13	4.6	24	3	US-08-884-029-7	Sequence 7, Appl
296	14	5.0	3170	3	US-09-344-520-1	Sequence 11, Appl	C 369	13	4.6	27	4	US-09-554-225-8	Sequence 8, Appl
297	14	5.0	3182	1	US-08-484-044-11	Sequence 11, Appl	C 370	13	4.6	34	3	US-09-282-996-11	Sequence 11, Appl
298	14	5.0	3344	1	US-08-426-236-1	Sequence 1, Appl	C 371	13	4.6	36	1	US-08-291-932A-525	Sequence 525, App
299	14	5.0	3394	1	US-08-159-784-4	Sequence 4, Appl	C 372	13	4.6	42	3	US-08-952-973-12	Sequence 12, Appl
300	14	5.0	3396	4	US-08-974-549A-642	Sequence 642, App	C 373	13	4.6	50	1	US-08-171-389-332	Sequence 332, App
301	14	5.0	3407	1	US-08-253-155A-7	Sequence 7, Appl	C 374	13	4.6	50	1	US-08-123-936-332	Sequence 332, App
302	14	5.0	3439	4	US-09-252-816A-6	Sequence 6, Appl	C 375	13	4.6	50	2	US-08-475-228A-332	Sequence 332, App
303	14	5.0	3729	1	US-08-554-612C-25	Sequence 25, Appl	C 376	13	4.6	50	3	US-08-482-080A-332	Sequence 332, App
304	14	5.0	3776	1	US-08-162-809-7	Sequence 7, Appl	C 377	13	4.6	50	5	PCT-US93-12388-332	Sequence 332, App
305	14	5.0	3808	2	US-08-916-917-3	Sequence 3, Appl	C 378	13	4.6	57	6	5221624-13	Patent No. 5221624
306	14	5.0	3808	2	US-08-972-631-3	Sequence 3, Appl	C 379	13	4.6	57	6	5221624-15	Patent No. 5221624
307	14	5.0	3808	2	US-08-972-629-3	Sequence 3, Appl	C 380	13	4.6	57	6	5221624-17	Patent No. 5221624
308	14	5.0	3808	2	US-08-972-630-3	Sequence 3, Appl	C 381	13	4.6	57	6	5221624-19	Patent No. 5221624
309	14	5.0	3808	2	US-08-672-211-3	Sequence 3, Appl	C 382	13	4.6	63	6	5221624-23	Patent No. 5221624
310	14	5.0	3808	1	US-09-225-170-3	Sequence 3, Appl	C 383	13	4.6	66	6	5221624-25	Patent No. 5221624
311	14	5.0	4085	1	US-08-072-574-7	Sequence 7, Appl	C 384	13	4.6	69	1	US-08-244-492A-8	Sequence 8, Appl
312	14	5.0	4085	1	US-08-486-270-7	Sequence 7, Appl	C 385	13	4.6	69	1	US-08-709-915-10	Sequence 10, Appl
313	14	5.0	4085	1	US-08-367-264-7	Sequence 7, Appl	C 386	13	4.6	73	3	US-08-952-973-21	Sequence 21, Appl
314	14	5.0	4134	2	US-08-817-090B-1	Sequence 1, Appl	C 387	13	4.6	75	5	PCT-US94-09653A-32	Sequence 32, Appl
315	14	5.0	4137	2	US-08-817-090B-3	Sequence 3, Appl	C 388	13	4.6	84	2	US-08-369-829A-15	Sequence 15, Appl
316	14	5.0	4181	1	US-08-072-574-9	Sequence 9, Appl	C 389	13	4.6	84	2	US-08-586-676E-20	Sequence 20, Appl
317	14	5.0	4181	1	US-08-486-270-9	Sequence 9, Appl	C 390	13	4.6	84	5	PCT-US94-09653A-33	Sequence 33, Appl
318	14	5.0	4181	1	US-08-367-264-9	Sequence 9, Appl	C 391	13	4.6	92	3	US-08-463-903-61	Sequence 61, Appl
319	14	5.0	4190	3	US-08-938-291A-2	Sequence 2, Appl	C 392	13	4.6	94	1	US-08-315-695-26	Sequence 26, Appl

C 393	13	4.6	104	1	US-08-073-807A-9	Sequence 9, Appl	C 466	13	4.6	603	2	US-08-924-838-6	Sequence 6, Appl
C 394	13	4.6	108	3	US-08-717-294-104	Sequence 104, App	C 467	13	4.6	613	1	US-08-998-416-112	Sequence 112, App
C 395	13	4.6	170	3	US-09-058-389A-13	Sequence 13, Appl	C 468	13	4.6	618	4	US-08-097-827-6	Sequence 6, Appl
C 396	13	4.6	180	1	US-08-443-640-2	Sequence 2, Appl	C 469	13	4.6	618	1	US-08-494-574-6	Sequence 6, Appl
C 397	13	4.6	196	4	US-09-221-298-76	Sequence 76, Appl	C 470	13	4.6	624	4	US-09-328-111-23	Sequence 213, App
C 398	13	4.6	206	4	US-09-476-124-11	Sequence 11, Appl	C 471	13	4.6	630	2	US-08-765-875-3	Sequence 3, Appl
C 399	13	4.6	207	4	US-09-476-124-8	Sequence 8, Appl	C 472	13	4.6	630	3	US-08-795-671-3	Sequence 38, Appl
C 400	13	4.6	208	4	US-09-476-124-43	Sequence 43, Appl	C 473	13	4.6	664	4	US-09-385-982-38	Sequence 279, App
C 401	13	4.6	239	4	US-08-944-483-3	Sequence 3, Appl	C 474	13	4.6	664	2	US-08-938-411-3	Sequence 3, Appl
C 402	13	4.6	256	4	US-09-251-372-3	Sequence 3, Appl	C 475	13	4.6	669	4	US-08-896-410-3	Sequence 342, App
C 403	13	4.6	259	2	US-08-493-197-3	Sequence 3, Appl	C 476	13	4.6	669	4	US-09-328-111-342	Sequence 159, Appl
C 404	13	4.6	259	2	PCT-US95-07844-3	Sequence 3, Appl	C 477	13	4.6	678	4	US-09-232-479-19	Sequence 55, Appl
C 405	13	4.6	273	4	US-08-882-164D-33	Sequence 33, Appl	C 478	13	4.6	711	4	US-08-793-701-55	Sequence 612, App
C 406	13	4.6	274	4	US-08-882-164D-34	Sequence 34, Appl	C 479	13	4.6	712	4	US-08-998-416-612	Sequence 701, App
C 407	13	4.6	277	2	US-08-634-797-38	Sequence 38, Appl	C 480	13	4.6	716	4	US-09-328-111-701	Sequence 892, App
C 408	13	4.6	285	3	US-09-026-343-25	Sequence 25, Appl	C 481	13	4.6	726	4	US-08-998-416-892	Sequence 2, Appl
C 409	13	4.6	294	3	US-09-060-756-409	Sequence 409, App	C 482	13	4.6	730	2	US-07-953-230A-2	Sequence 9, Appl
C 410	13	4.6	297	3	US-09-026-343-9	Sequence 9, Appl	C 483	13	4.6	750	2	US-08-627-151A-9	Sequence 85, Appl
C 411	13	4.6	297	3	US-09-026-343-23	Sequence 23, Appl	C 484	13	4.6	751	1	US-08-599-252-85	Sequence 32, Appl
C 412	13	4.6	319	2	US-08-244-537-11	Sequence 11, Appl	C 485	13	4.6	751	4	US-08-896-164-32	Sequence 85, Appl
C 413	13	4.6	331	2	US-09-060-756-310	Sequence 310, App	C 486	13	4.6	751	5	PCT-US96-06352-85	Sequence 85, Appl
C 414	13	4.6	373	2	US-08-791-849A-10	Sequence 10, Appl	C 487	13	4.6	751	5	PCT-US96-06583-85	Sequence 9, Appl
C 415	13	4.6	390	1	US-08-482-882-79	Sequence 79, Appl	C 488	13	4.6	752	4	US-09-378-255-9	Sequence 85, Appl
C 416	13	4.6	390	1	US-08-483-389-79	Sequence 79, Appl	C 489	13	4.6	752	4	US-09-715-336-9	Sequence 9, Appl
C 417	13	4.6	390	2	US-08-487-113D-79	Sequence 79, Appl	C 490	13	4.6	757	1	US-08-128-587C-6	Sequence 6, Appl
C 418	13	4.6	390	2	US-08-473-503-79	Sequence 79, Appl	C 491	13	4.6	766	4	US-08-998-416-395	Sequence 395, App
C 419	13	4.6	390	2	US-08-483-932-79	Sequence 79, Appl	C 492	13	4.6	801	2	US-08-934-481-3	Sequence 3, Appl
C 420	13	4.6	390	2	US-08-720-420A-79	Sequence 79, Appl	C 493	13	4.6	801	4	US-09-227-531-111	Sequence 111, App
C 421	13	4.6	390	3	US-08-714-017-79	Sequence 79, Appl	C 494	13	4.6	804	2	US-08-835-039A-10	Sequence 10, Appl
C 422	13	4.6	390	3	US-08-475-680-79	Sequence 79, Appl	C 495	13	4.6	804	3	US-09-157-349-10	Sequence 10, Appl
C 423	13	4.6	402	1	US-08-435-454-1	Sequence 1, Appl	C 496	13	4.6	826	4	US-08-998-416-561	Sequence 561, App
C 424	13	4.6	402	1	US-08-652-972A-1	Sequence 1, Appl	C 497	13	4.6	836	2	US-08-790-137-2	Sequence 2, Appl
C 425	13	4.6	402	2	US-08-919-145-3	Sequence 3, Appl	C 498	13	4.6	846	2	US-08-463-903-1	Sequence 43, Appl
C 426	13	4.6	402	2	US-08-870-126-1	Sequence 1, Appl	C 499	13	4.6	846	2	US-08-619-542B-43	Sequence 16, Appl
C 427	13	4.6	402	5	US-09-344-889-3	Sequence 3, Appl	C 500	13	4.6	856	3	US-08-463-903-16	Sequence 16, Appl
C 428	13	4.6	402	5	PCT-US96-06231A-1	Sequence 1, Appl	C 501	13	4.6	857	4	US-09-318-448-16	Sequence 10, Appl
C 429	13	4.6	413	4	US-09-342-653-1	Sequence 1, Appl	C 502	13	4.6	872	1	US-08-374-983A-10	Sequence 14, Appl
C 430	13	4.6	416	2	US-08-890-980-9	Sequence 9, Appl	C 503	13	4.6	876	4	US-08-446-935-2	Sequence 2, Appl
C 431	13	4.6	416	3	US-08-890-979-9	Sequence 9, Appl	C 504	13	4.6	877	3	US-08-463-903-14	Sequence 14, Appl
C 432	13	4.6	416	3	US-09-032-894-9	Sequence 9, Appl	C 505	13	4.6	890	1	US-08-592-126-78	Sequence 78, Appl
C 433	13	4.6	416	3	US-09-031-626-9	Sequence 9, Appl	C 506	13	4.6	903	4	US-09-457-068-5	Sequence 5, Appl
C 434	13	4.6	418	4	US-08-467-504-15	Sequence 15, Appl	C 507	13	4.6	908	4	US-08-718-388-1	Sequence 1, Appl
C 435	13	4.6	441	1	US-08-259-372A-7	Sequence 7, Appl	C 508	13	4.6	909	2	US-08-933-750C-84	Sequence 84, Appl
C 436	13	4.6	441	1	US-08-468-671-7	Sequence 7, Appl	C 509	13	4.6	909	3	US-09-234-613-84	Sequence 84, Appl
C 437	13	4.6	449	1	US-08-461-284-3	Sequence 3, Appl	C 510	13	4.6	929	2	US-08-586-676E-9	Sequence 9, Appl
C 438	13	4.6	449	1	US-08-462-939-3	Sequence 3, Appl	C 511	13	4.6	947	1	US-08-096-623A-16	Sequence 16, Appl
C 439	13	4.6	449	1	US-08-253-877C-3	Sequence 3, Appl	C 512	13	4.6	947	1	US-08-096-623A-17	Sequence 17, Appl
C 440	13	4.6	449	2	US-08-452-164A-3	Sequence 3, Appl	C 513	13	4.6	965	2	US-09-141-135-1	Sequence 1, Appl
C 441	13	4.6	456	4	US-09-040-984-79	Sequence 79, Appl	C 514	13	4.6	984	1	US-08-611-107-11	Sequence 11, Appl
C 442	13	4.6	461	4	US-08-940-424-11	Sequence 11, Appl	C 515	13	4.6	984	2	US-08-422-560A-11	Sequence 11, Appl
C 443	13	4.6	471	1	US-08-616-368A-38	Sequence 38, Appl	C 516	13	4.6	984	4	US-08-468-793-11	Sequence 25, Appl
C 444	13	4.6	471	3	US-09-054-298-38	Sequence 38, Appl	C 517	13	4.6	997	1	US-08-480-784-25	Sequence 25, Appl
C 445	13	4.6	471	3	US-08-818-655-38	Sequence 38, Appl	C 518	13	4.6	997	1	US-08-483-553-25	Sequence 25, Appl
C 446	13	4.6	475	2	US-08-623-906A-4	Sequence 4, Appl	C 519	13	4.6	997	1	US-08-487-002-25	Sequence 25, Appl
C 447	13	4.6	486	1	US-08-324-243-36	Sequence 36, Appl	C 520	13	4.6	997	1	US-08-483-554B-25	Sequence 25, Appl
C 448	13	4.6	486	1	US-08-532-390-36	Sequence 36, Appl	C 521	13	4.6	997	1	US-08-488-011B-25	Sequence 25, Appl
C 449	13	4.6	486	5	US-08-717-294-36	Sequence 36, Appl	C 522	13	4.6	997	4	US-08-850-727-95	Sequence 25, Appl
C 450	13	4.6	486	5	PCT-US95-11511-36	Sequence 36, Appl	C 523	13	4.6	997	5	PCT-US95-10202-25	Sequence 25, Appl
C 451	13	4.6	525	6	5422248-3	Patent No. 5422248	C 524	13	4.6	997	5	PCT-US95-10203-25	Sequence 25, Appl
C 452	13	4.6	534	4	US-09-030-607-187	Sequence 187, App	C 525	13	4.6	997	5	PCT-US95-10220-25	Sequence 25, Appl
C 453	13	4.6	546	4	US-08-971-090-1	Sequence 1, Appl	C 526	13	4.6	1003	3	US-08-961-083-109	Sequence 109, App
C 454	13	4.6	559	1	US-08-469-667-19	Sequence 19, Appl	C 527	13	4.6	1006	4	US-09-058-368-4	Sequence 4, Appl
C 455	13	4.6	559	5	PCT-US95-07289-19	Sequence 19, Appl	C 528	13	4.6	1027	4	US-08-991-789A-199	Sequence 199, App
C 456	13	4.6	574	4	US-09-189-060B-18	Sequence 18, Appl	C 529	13	4.6	1035	2	US-08-690-493-2	Sequence 2, Appl
C 457	13	4.6	575	3	US-08-726-807B-6	Sequence 6, Appl	C 530	13	4.6	1037	2	US-08-824-405-1	Sequence 1, Appl
C 458	13	4.6	575	3	US-09-258-367-6	Sequence 6, Appl	C 531	13	4.6	1041	2	US-08-602-359A-24	Sequence 24, Appl
C 459	13	4.6	575	4	US-09-546-550-6	Sequence 6, Appl	C 532	13	4.6	1083	3	US-08-331-004A-1	Sequence 1, Appl
C 460	13	4.6	575	4	US-09-431-414-6	Sequence 6, Appl	C 533	13	4.6	1083	5	PCT-US95-13937A-1	Sequence 1, Appl
C 461	13	4.6	575	4	US-09-225-670-6	Sequence 6, Appl	C 534	13	4.6	1102	3	US-08-726-807B-8	Sequence 8, Appl
C 462	13	4.6	576	1	US-08-086-428B-37	Sequence 37, Appl	C 535	13	4.6	1102	3	US-09-258-367-8	Sequence 8, Appl
C 463	13	4.6	576	2	US-08-468-570-37	Sequence 37, Appl	C 536	13	4.6	1102	4	US-09-546-550-8	Sequence 8, Appl
C 464	13	4.6	576	2	US-08-290-665A-37	Sequence 37, Appl	C 537	13	4.6	1102	4	US-09-431-414-8	Sequence 8, Appl
C 465	13	4.6	576	5	PCT-US95-10398-37	Sequence 37, Appl	C 538	13	4.6	1102	4	US-09-225-670-8	Sequence 8, Appl

539	13	4.6	1105	2	US-08-466-103A-15	Sequence 15, Appl	C 612	13	4.6	1497	4	US-09-232-468A-23	Sequence 23, Appl
540	13	4.6	1119	1	US-08-416-756A-1	Sequence 1, Appl1	613	13	4.6	1515	1	US-07-745-206A-22	Sequence 22, Appl
541	13	4.6	1128	2	US-08-446-875-7	Sequence 7, Appl1	C 614	13	4.6	1515	1	US-07-745-206A-22	Sequence 22, Appl
542	13	4.6	1128	2	US-08-102-385G-17	Sequence 7, Appl1	615	13	4.6	1515	2	US-08-311-363-22	Sequence 22, Appl
543	13	4.6	1128	2	US-08-562-114B-14	Sequence 14, Appl	C 616	13	4.6	1515	2	US-08-311-363-22	Sequence 22, Appl
544	13	4.6	1128	4	US-08-729-594A-14	Sequence 14, Appl	617	13	4.6	1526	1	US-08-455-543A-38	Sequence 38, Appl
545	13	4.6	1129	1	US-08-470-261-1	Sequence 1, Appl1	618	13	4.6	1526	2	US-08-193-078B-27	Sequence 27, Appl
546	13	4.6	1129	3	US-08-916-989B-1	Sequence 1, Appl1	619	13	4.6	1526	2	US-08-223-305C-38	Sequence 38, Appl
547	13	4.6	1129	5	PCT-US94-13187-1	Sequence 7, Appl1	620	13	4.6	1526	3	US-08-149-097B-30	Sequence 30, Appl
548	13	4.6	1166	4	US-08-944-483-7	Sequence 2, Appl1	621	13	4.6	1526	3	US-08-949-386-34	Sequence 34, Appl
549	13	4.6	1181	2	US-08-632-598-2	Sequence 2, Appl1	622	13	4.6	1526	3	US-08-450-562-34	Sequence 34, Appl
550	13	4.6	1181	4	US-09-231-240-2	Sequence 2, Appl1	C 623	13	4.6	1527	2	US-08-850-293-4	Sequence 4, Appl1
551	13	4.6	1189	4	US-09-497-779A-7	Sequence 7, Appl1	624	13	4.6	1546	1	US-07-745-206A-18	Sequence 18, Appl1
552	13	4.6	1192	4	US-08-944-483-8	Sequence 8, Appl1	625	13	4.6	1546	1	US-08-455-543A-9	Sequence 9, Appl1
553	13	4.6	1198	4	US-08-095-726-5	Sequence 5, Appl1	626	13	4.6	1546	2	US-08-193-078B-9	Sequence 9, Appl1
554	13	4.6	1198	1	US-08-096-043-5	Sequence 5, Appl1	627	13	4.6	1546	2	US-08-223-305C-39	Sequence 9, Appl1
555	13	4.6	1198	1	US-08-093-577-5	Sequence 5, Appl1	628	13	4.6	1546	2	US-08-149-097B-9	Sequence 9, Appl1
556	13	4.6	1198	1	US-08-096-623A-5	Sequence 5, Appl1	629	13	4.6	1546	2	US-08-311-363-18	Sequence 18, Appl1
557	13	4.6	1200	2	US-08-690-493-3	Sequence 3, Appl1	630	13	4.6	1546	3	US-08-949-386-9	Sequence 9, Appl1
558	13	4.6	1212	2	US-09-092-770-2	Sequence 2, Appl1	631	13	4.6	1546	3	US-08-450-562-9	Sequence 9, Appl1
559	13	4.6	1212	4	US-09-222-851-2	Sequence 2, Appl1	C 632	13	4.6	1547	2	US-08-902-853-2	Sequence 2, Appl1
560	13	4.6	1218	1	US-08-673-269-1	Sequence 1, Appl1	633	13	4.6	1566	2	US-08-145-658B-22	Sequence 22, Appl1
561	13	4.6	1223	4	US-08-829-110-4	Sequence 4, Appl1	634	13	4.6	1566	2	US-08-145-658B-22	Sequence 22, Appl1
562	13	4.6	1229	4	US-08-940-424-10	Sequence 10, Appl	635	13	4.6	1571	2	US-08-145-658B-21	Sequence 21, Appl
563	13	4.6	1230	4	US-08-940-424-10	Sequence 10, Appl	636	13	4.6	1581	2	US-08-313-553-8	Sequence 21, Appl
564	13	4.6	1248	3	US-08-897-340-5	Sequence 5, Appl1	637	13	4.6	1590	1	US-08-767-993-8	Sequence 8, Appl1
565	13	4.6	1248	3	US-09-252-329-5	Sequence 5, Appl1	C 638	13	4.6	1590	2	US-08-836-442-2	Sequence 2, Appl1
566	13	4.6	1251	4	US-08-261-206A-75	Sequence 75, Appl	C 639	13	4.6	1593	2	US-08-524-828-2	Sequence 2, Appl1
567	13	4.6	1257	4	US-09-554-225-6	Sequence 2, Appl1	C 640	13	4.6	1593	2	US-08-975-114A-2	Sequence 2, Appl1
568	13	4.6	1260	3	US-08-949-202-1	Sequence 1, Appl1	C 641	13	4.6	1593	3	US-08-849-281A-2	Sequence 2, Appl1
569	13	4.6	1260	4	US-09-418-175-1	Sequence 1, Appl1	642	13	4.6	1594	1	US-08-248-628A-1	Sequence 1, Appl1
570	13	4.6	1267	4	US-09-422-869-15	Sequence 15, Appl	643	13	4.6	1596	1	US-08-513-841-3	Sequence 3, Appl1
571	13	4.6	1278	4	US-08-940-424-12	Sequence 12, Appl	644	13	4.6	1596	2	US-08-696-834-3	Sequence 3, Appl1
572	13	4.6	1279	1	US-08-146-010A-4	Sequence 4, Appl1	645	13	4.6	1596	2	US-08-942-673-3	Sequence 3, Appl1
573	13	4.6	1279	3	US-08-674-168-9	Sequence 9, Appl1	646	13	4.6	1596	4	US-09-118-317-3	Sequence 3, Appl1
574	13	4.6	1279	3	US-09-248-335-25	Sequence 25, Appl	C 647	13	4.6	1602	4	US-09-360-197-9	Sequence 9, Appl1
575	13	4.6	1296	1	US-08-578-709-12	Sequence 12, Appl	648	13	4.6	1612	1	US-08-343-733A-1	Sequence 1, Appl1
576	13	4.6	1305	2	US-08-896-410-37	Sequence 37, Appl	649	13	4.6	1630	3	US-08-996-139-10	Sequence 10, Appl
577	13	4.6	1309	6	5221624-4	Patent No. 5221624	650	13	4.6	1630	4	US-08-995-659-10	Sequence 10, Appl
578	13	4.6	1311	1	US-08-242-098-38	Sequence 38, Appl	651	13	4.6	1648	4	US-09-215-649A-10	Sequence 10, Appl
579	13	4.6	1314	3	US-09-025-059-2	Sequence 2, Appl1	C 652	13	4.6	1666	4	US-08-732-749-3	Sequence 3, Appl1
580	13	4.6	1317	1	US-08-097-827-10	Sequence 10, Appl	C 653	13	4.6	1666	4	US-09-149-922-4	Sequence 4, Appl1
581	13	4.6	1317	1	US-08-494-574-10	Sequence 10, Appl	654	13	4.6	1681	1	US-08-455-543A-37	Sequence 37, Appl
582	13	4.6	1334	4	US-08-940-424-13	Sequence 13, Appl	655	13	4.6	1681	2	US-08-193-078B-26	Sequence 26, Appl
583	13	4.6	1336	4	US-08-718-388-2	Sequence 2, Appl1	656	13	4.6	1681	2	US-08-223-305C-37	Sequence 37, Appl
584	13	4.6	1346	3	US-08-949-202-3	Sequence 3, Appl1	657	13	4.6	1681	2	US-08-149-097B-29	Sequence 29, Appl
585	13	4.6	1346	4	US-09-418-175-3	Sequence 3, Appl1	658	13	4.6	1681	2	US-08-581-527-4	Sequence 4, Appl1
586	13	4.6	1350	1	US-08-665-435A-1	Sequence 1, Appl1	659	13	4.6	1681	3	US-08-949-386-33	Sequence 33, Appl
587	13	4.6	1350	2	US-08-665-435A-3	Sequence 3, Appl1	660	13	4.6	1681	3	US-08-450-562-33	Sequence 33, Appl
588	13	4.6	1350	2	US-08-843-309-3	Sequence 3, Appl1	661	13	4.6	1681	5	PCT-US94-07754-4	Sequence 4, Appl1
589	13	4.6	1353	3	US-08-843-309-1	Sequence 9, Appl1	662	13	4.6	1683	1	US-08-164-614A-5	Sequence 5, Appl1
590	13	4.6	1370	4	US-09-111-470-9	Sequence 9, Appl1	663	13	4.6	1683	2	US-08-824-405-5	Sequence 5, Appl1
591	13	4.6	1388	4	US-09-008-271A-22	Sequence 22, Appl	664	13	4.6	1683	2	US-08-456-489B-5	Sequence 5, Appl1
592	13	4.6	1393	1	US-08-455-543A-39	Sequence 39, Appl	665	13	4.6	1683	3	PCT-US93-01720-5	Sequence 5, Appl1
593	13	4.6	1393	2	US-08-193-078B-28	Sequence 28, Appl	666	13	4.6	1701	1	US-07-939-501A-11	Sequence 11, Appl
594	13	4.6	1393	2	US-08-223-305C-39	Sequence 39, Appl	667	13	4.6	1702	1	US-08-455-543A-19	Sequence 19, Appl
595	13	4.6	1393	2	US-08-149-097B-31	Sequence 31, Appl	668	13	4.6	1702	2	US-08-223-305C-19	Sequence 19, Appl
596	13	4.6	1393	3	US-08-949-386-35	Sequence 35, Appl	669	13	4.6	1726	2	US-08-467-023-133	Sequence 133, Appl
597	13	4.6	1393	3	US-08-450-562-35	Sequence 35, Appl	669	13	4.6	1750	3	US-08-859-167-1	Sequence 1, Appl1
598	13	4.6	1413	4	US-09-342-459-1	Sequence 1, Appl1	C 670	13	4.6	1750	3	US-08-859-167-1	Sequence 1, Appl1
599	13	4.6	1432	5	PCT-US94-12912-4	Sequence 4, Appl1	C 671	13	4.6	1750	3	US-09-109-273-1	Sequence 1, Appl1
600	13	4.6	1450	2	US-09-008-962-2	Sequence 2, Appl1	C 672	13	4.6	1755	4	US-09-276-99-15	Sequence 15, Appl
601	13	4.6	1450	2	US-08-675-507-2	Sequence 2, Appl1	673	13	4.6	1755	4	US-08-543-246B-8	Sequence 8, Appl1
602	13	4.6	1450	2	US-09-213-205-2	Sequence 2, Appl1	674	13	4.6	1772	1	US-08-336-257A-2	Sequence 2, Appl1
603	13	4.6	1453	6	5194600-1	Patent No. 5194600	675	13	4.6	1772	6	5386025-3	Sequence 3, Appl1
604	13	4.6	1458	4	US-09-111-470-3	Sequence 3, Appl	676	13	4.6	1773	4	US-08-943-731-215	Sequence 215, Appl
605	13	4.6	1468	3	US-09-079-981-2	Sequence 2, Appl1	677	13	4.6	1782	4	US-08-153-799-15	Sequence 15, Appl
606	13	4.6	1476	3	US-08-713-118-5	Sequence 5, Appl1	678	13	4.6	1782	4	US-08-769-746-1	Sequence 1, Appl1
607	13	4.6	1476	3	US-09-452-007-5	Sequence 5, Appl1	C 679	13	4.6	1784	3	US-08-991-426-1	Sequence 1, Appl1
608	13	4.6	1497	2	US-08-820-170A-41	Sequence 41, Appl	C 680	13	4.6	1784	3	US-09-143-470-1	Sequence 1, Appl1
609	13	4.6	1497	3	US-09-055-699-41	Sequence 41, Appl	681	13	4.6	1795	5	US-08-330-517-1	Sequence 1, Appl1
610	13	4.6	1497	4	US-09-273-565-41	Sequence 41, Appl	682	13	4.6	1801	5	PCT-US95-02455-1	Sequence 1, Appl1
611	13	4.6	1497	4	US-09-232-468A-17	Sequence 17, Appl	683	13	4.6	1805	1	US-08-343-733A-3	Sequence 3, Appl1
							684	13	4.6	1830	1	US-07-691-079C-3	Sequence 3, Appl1

685	13	4.6	1830	1	US-08-343-733A-2	Sequence 2, Appl	c 758	13	4.6	1945	4	US-09-265-315-77	Sequence 77, Appl
686	13	4.6	1830	1	US-08-433-037-3	Sequence 3, Appl	c 759	13	4.6	1945	4	US-09-265-315-77	Sequence 77, Appl
687	13	4.6	1830	2	US-08-702-572-1	Sequence 10, Appl	c 760	13	4.6	1945	4	US-09-266-417-77	Sequence 77, Appl
688	13	4.6	1851	1	US-08-455-543A-10	Sequence 10, Appl	c 761	13	4.6	1947	2	US-08-164-614A-4	Sequence 4, Appl
689	13	4.6	1851	2	US-08-193-078B-10	Sequence 10, Appl	c 762	13	4.6	1947	2	US-08-456-489B-6	Sequence 4, Appl
690	13	4.6	1851	2	US-08-223-305C-10	Sequence 10, Appl	c 763	13	4.6	1947	5	PCT-US93-01720-4	Sequence 4, Appl
691	13	4.6	1851	2	US-08-149-097D-10	Sequence 10, Appl	c 764	13	4.6	1956	3	US-08-313-553-6	Sequence 6, Appl
692	13	4.6	1851	2	US-08-949-386-10	Sequence 10, Appl	c 765	13	4.6	1956	3	US-08-767-993-6	Sequence 6, Appl
693	13	4.6	1851	3	US-08-450-562-10	Sequence 10, Appl	c 766	13	4.6	1959	4	US-09-061-764A-4	Sequence 4, Appl
694	13	4.6	1852	2	US-08-797-689-1	Sequence 1, Appl	c 767	13	4.6	1977	1	US-08-578-709-14	Sequence 14, Appl
695	13	4.6	1870	3	US-09-295-186-9	Sequence 9, Appl	c 768	13	4.6	1979	4	US-08-552-309-19	Sequence 19, Appl
c 696	13	4.6	1870	4	US-09-476-124-9	Sequence 9, Appl	c 769	13	4.6	1997	2	US-08-164-614A-6	Sequence 6, Appl
c 697	13	4.6	1876	4	US-09-082-092-5	Sequence 5, Appl	c 770	13	4.6	1997	2	US-08-456-489B-6	Sequence 6, Appl
c 698	13	4.6	1881	4	US-09-029-348-20	Sequence 20, Appl	c 771	13	4.6	1997	5	US-08-963-630-1	Sequence 1, Appl
699	13	4.6	1882	3	US-08-501-572-4	Sequence 4, Appl	c 772	13	4.6	1997	5	PCT-US93-01720-6	Sequence 6, Appl
700	13	4.6	1882	3	US-09-040-444-4	Sequence 4, Appl	c 773	13	4.6	2005	1	US-07-916-901-5	Sequence 5, Appl
c 701	13	4.6	1890	3	US-09-289-466-2	Sequence 2, Appl	c 774	13	4.6	2006	2	US-08-149-097D-19	Sequence 19, Appl
c 702	13	4.6	1891	3	US-09-289-466-1	Sequence 1, Appl	c 775	13	4.6	2033	1	US-08-148-910-14	Sequence 14, Appl
c 703	13	4.6	1892	4	US-09-554-225-3	Sequence 3, Appl	c 776	13	4.6	2036	1	US-08-448-937A-14	Sequence 14, Appl
c 704	13	4.6	1896	4	US-09-232-220-31	Sequence 31, Appl	c 777	13	4.6	2042	2	US-08-923-457A-17	Sequence 17, Appl
c 705	13	4.6	1896	4	US-09-232-197-31	Sequence 31, Appl	c 778	13	4.6	2042	2	US-08-616-844-5	Sequence 5, Appl
c 706	13	4.6	1905	1	US-08-164-614A-2	Sequence 2, Appl	c 779	13	4.6	2042	2	US-08-599-654-5	Sequence 5, Appl
c 707	13	4.6	1905	2	US-08-456-489B-2	Sequence 2, Appl	c 780	13	4.6	2042	3	US-08-944-868A-5	Sequence 5, Appl
c 708	13	4.6	1905	5	PCT-US93-01720-2	Sequence 2, Appl	c 781	13	4.6	2042	3	US-08-944-423A-5	Sequence 5, Appl
709	13	4.6	1917	4	US-08-808-346-1	Sequence 1, Appl	c 782	13	4.6	2042	3	US-08-944-423A-5	Sequence 5, Appl
710	13	4.6	1919	4	US-08-975-762-40	Sequence 40, Appl	c 783	13	4.6	2042	3	US-08-944-423A-5	Sequence 5, Appl
711	13	4.6	1919	4	US-09-295-028-40	Sequence 40, Appl	c 784	13	4.6	2045	2	US-08-795-088A-1	Sequence 1, Appl
712	13	4.6	1919	4	US-09-106-582-40	Sequence 40, Appl	c 785	13	4.6	2057	2	US-08-836-442-1	Sequence 1, Appl
c 713	13	4.6	1925	1	US-08-435-454-3	Sequence 3, Appl	c 786	13	4.6	2057	2	US-08-820-170A-42	Sequence 42, Appl
c 714	13	4.6	1925	1	US-08-652-972A-3	Sequence 3, Appl	c 787	13	4.6	2057	4	US-09-055-699-42	Sequence 42, Appl
c 715	13	4.6	1925	2	US-08-919-145-5	Sequence 5, Appl	c 788	13	4.6	2061	4	US-09-273-565-42	Sequence 42, Appl
c 716	13	4.6	1925	3	US-08-870-126-3	Sequence 3, Appl	c 789	13	4.6	2069	1	US-09-061-764A-17	Sequence 17, Appl
c 717	13	4.6	1925	5	US-09-344-889-5	Sequence 5, Appl	c 790	13	4.6	2089	5	US-08-552-142A-1	Sequence 1, Appl
c 718	13	4.6	1925	5	PCT-US96-06231A-3	Sequence 3, Appl	c 791	13	4.6	2089	5	US-08-910-973-1	Sequence 1, Appl
719	13	4.6	1926	1	US-07-901-703-12	Sequence 12, Appl	c 792	13	4.6	2098	2	US-08-937-931-5	Sequence 5, Appl
720	13	4.6	1926	1	US-08-147-023-26	Sequence 26, Appl	c 793	13	4.6	2098	4	US-09-283-502-5	Sequence 5, Appl
721	13	4.6	1926	1	US-08-278-729A-22	Sequence 22, Appl	c 794	13	4.6	2101	6	US-08-106-761-1	Sequence 1, Appl
722	13	4.6	1926	1	US-08-480-528A-9	Sequence 9, Appl	c 795	13	4.6	2124	6	5428012-1	Patent No. 5428012
723	13	4.6	1926	1	US-08-479-666-9	Sequence 9, Appl	c 796	13	4.6	2124	6	5451506-1	Patent No. 5451506
724	13	4.6	1926	1	US-08-155-343A-22	Sequence 22, Appl	c 797	13	4.6	2139	3	US-09-026-343-1	Sequence 1, Appl
725	13	4.6	1926	1	US-08-406-672-22	Sequence 22, Appl	c 798	13	4.6	2144	3	US-08-149-097D-20	Sequence 20, Appl
726	13	4.6	1926	1	US-08-643-563A-22	Sequence 22, Appl	c 799	13	4.6	2144	3	US-08-949-386-20	Sequence 20, Appl
727	13	4.6	1926	1	US-08-447-570-26	Sequence 26, Appl	c 800	13	4.6	2150	2	US-08-450-562-20	Sequence 20, Appl
728	13	4.6	1926	1	US-08-643-763A-22	Sequence 22, Appl	c 801	13	4.6	2150	1	US-08-580-680-2	Sequence 2, Appl
729	13	4.6	1926	1	US-08-462-623-22	Sequence 22, Appl	c 802	13	4.6	2150	1	US-08-480-156A-2	Sequence 2, Appl
730	13	4.6	1926	1	US-08-451-953A-22	Sequence 22, Appl	c 803	13	4.6	2150	1	US-08-354-961-2	Sequence 2, Appl
731	13	4.6	1926	2	US-08-459-346-7	Sequence 7, Appl	c 804	13	4.6	2150	2	US-08-581-094-2	Sequence 2, Appl
732	13	4.6	1926	2	US-08-445-468A-22	Sequence 22, Appl	c 805	13	4.6	2150	2	US-08-580-665-2	Sequence 2, Appl
733	13	4.6	1926	2	US-08-901-200A-9	Sequence 9, Appl	c 806	13	4.6	2150	2	US-08-581-142-2	Sequence 2, Appl
734	13	4.6	1926	2	US-08-449-700-26	Sequence 26, Appl	c 807	13	4.6	2150	5	PCT-US94-05442A-2	Sequence 2, Appl
735	13	4.6	1926	2	US-08-449-699A-26	Sequence 26, Appl	c 808	13	4.6	2150	5	PCT-US94-05442A-2	Sequence 2, Appl
736	13	4.6	1926	2	US-08-461-397A-22	Sequence 22, Appl	c 809	13	4.6	2151	2	US-08-944-449-3	Sequence 3, Appl
737	13	4.6	1926	2	US-08-912-088-22	Sequence 22, Appl	c 810	13	4.6	2153	3	US-08-949-386-19	Sequence 19, Appl
738	13	4.6	1926	3	US-08-278-730A-22	Sequence 22, Appl	c 811	13	4.6	2153	3	US-08-450-562-19	Sequence 19, Appl
739	13	4.6	1926	3	US-08-889-419-7	Sequence 7, Appl	c 812	13	4.6	2164	4	US-08-450-562-19	Sequence 19, Appl
740	13	4.6	1926	3	US-08-445-467-22	Sequence 22, Appl	c 813	13	4.6	2164	4	US-09-318-793A-3	Sequence 3, Appl
741	13	4.6	1926	3	US-08-480-515A-22	Sequence 22, Appl	c 814	13	4.6	2191	4	US-08-989-362-1	Sequence 1, Appl
742	13	4.6	1926	3	US-09-219-391-9	Sequence 9, Appl	c 815	13	4.6	2193	4	US-08-934-861-1	Sequence 1, Appl
743	13	4.6	1926	5	PCT-US92-01968-22	Sequence 22, Appl	c 816	13	4.6	2200	2	US-08-462-481-3	Sequence 3, Appl
744	13	4.6	1926	5	PCT-US93-05446-12	Sequence 12, Appl	c 817	13	4.6	2200	2	US-08-436-771-5	Sequence 5, Appl
745	13	4.6	1926	5	PCT-US93-07189-7	Sequence 7, Appl	c 818	13	4.6	2200	2	US-08-434-998-5	Sequence 5, Appl
746	13	4.6	1926	5	PCT-US93-07190-22	Sequence 22, Appl	c 819	13	4.6	2200	2	US-08-487-797-5	Sequence 5, Appl
747	13	4.6	1926	5	PCT-US93-07231-22	Sequence 22, Appl	c 820	13	4.6	2200	2	US-08-701-005A-3	Sequence 3, Appl
748	13	4.6	1926	5	PCT-US93-08742-22	Sequence 22, Appl	c 821	13	4.6	2200	2	US-08-479-895-3	Sequence 3, Appl
749	13	4.6	1926	5	PCT-US93-08808-22	Sequence 22, Appl	c 822	13	4.6	2200	2	US-08-943-956A-3	Sequence 3, Appl
750	13	4.6	1926	5	PCT-US93-08885-22	Sequence 22, Appl	c 823	13	4.6	2200	5	PCT-US95-02058-5	Sequence 5, Appl
751	13	4.6	1926	5	PCT-US93-10520-9	Sequence 9, Appl	c 824	13	4.6	2204	1	US-08-221-817-12	Sequence 12, Appl
752	13	4.6	1929	1	US-07-841-646-26	Sequence 26, Appl	c 825	13	4.6	2204	1	US-08-454-439-12	Sequence 12, Appl
c 753	13	4.6	1929	2	US-09-016-000-10	Sequence 10, Appl	c 826	13	4.6	2204	5	PCT-US94-10487-12	Sequence 12, Appl
c 754	13	4.6	1929	2	PCT-US91-07635-3	Sequence 3, Appl	c 827	13	4.6	2205	3	US-08-888-077A-11	Sequence 11, Appl
c 755	13	4.6	1933	3	US-08-627-907A-1	Sequence 1, Appl	c 828	13	4.6	2205	4	US-08-687-590-59	Sequence 59, Appl
c 756	13	4.6	1940	3	US-08-732-749-1	Sequence 1, Appl	c 829	13	4.6	2206	4	US-08-221-817-10	Sequence 10, Appl
c 757	13	4.6	1945	3	US-08-714-918-77	Sequence 77, Appl	c 830	13	4.6	2206	1	US-08-454-439-10	Sequence 10, Appl

C 831	13	4.6	2206	5	PCT-US94-10487-10	Sequence 10, Appl	C 904	13	4.6	2609	4	US-09-141-212-7	Sequence 7, Appl
C 832	13	4.6	2231	4	US-09-008-271A-24	Sequence 24, Appl	C 905	13	4.6	2609	4	US-09-561-138-7	Sequence 7, Appl
C 833	13	4.6	2247	2	US-08-524-828-1	Sequence 1, Appl	C 906	13	4.6	2612	4	US-09-042-785A-3	Sequence 3, Appl
C 834	13	4.6	2247	4	US-08-975-114A-1	Sequence 1, Appl	C 907	13	4.6	2630	2	US-08-890-980-1	Sequence 1, Appl
C 835	13	4.6	2250	3	US-08-763-938-1	Sequence 1, Appl	C 908	13	4.6	2630	3	US-08-890-979-1	Sequence 1, Appl
C 836	13	4.6	2251	4	US-08-991-677-11	Sequence 11, Appl	C 909	13	4.6	2630	3	US-09-032-894-1	Sequence 1, Appl
C 837	13	4.6	2256	1	US-08-181-271A-2	Sequence 2, Appl	C 910	13	4.6	2630	4	US-09-031-626-1	Sequence 1, Appl
C 838	13	4.6	2256	1	US-08-449-315-2	Sequence 2, Appl	C 911	13	4.6	2661	1	US-08-592-126-70	Sequence 70, Appl
C 839	13	4.6	2256	1	US-08-444-803-2	Sequence 2, Appl	C 912	13	4.6	2663	2	US-08-653-740-4	Sequence 4, Appl
C 840	13	4.6	2256	1	US-08-449-043-2	Sequence 2, Appl	C 913	13	4.6	2663	2	US-09-073-594-4	Sequence 4, Appl
C 841	13	4.6	2256	1	US-08-456-265A-2	Sequence 2, Appl	C 914	13	4.6	2663	3	US-09-275-925-4	Sequence 4, Appl
C 842	13	4.6	2256	1	US-08-455-416-2	Sequence 2, Appl	C 915	13	4.6	2669	4	US-09-141-212-9	Sequence 9, Appl
C 843	13	4.6	2256	1	US-08-455-244-2	Sequence 2, Appl	C 916	13	4.6	2669	4	US-09-561-138-9	Sequence 9, Appl
C 844	13	4.6	2256	1	US-08-454-876-2	Sequence 2, Appl	C 917	13	4.6	2672	3	US-09-214-564A-5	Sequence 5, Appl
C 845	13	4.6	2256	2	US-08-457-364-2	Sequence 2, Appl	C 918	13	4.6	2677	4	US-08-882-164D-36	Sequence 36, Appl
C 846	13	4.6	2256	2	US-08-456-262-2	Sequence 2, Appl	C 919	13	4.6	2688	1	US-08-088-633-3	Sequence 3, Appl
C 847	13	4.6	2256	2	US-08-456-240-2	Sequence 2, Appl	C 920	13	4.6	2688	1	US-08-245-756-3	Sequence 3, Appl
C 848	13	4.6	2256	2	US-08-455-736-2	Sequence 2, Appl	C 921	13	4.6	2688	1	US-08-441-750-3	Sequence 3, Appl
C 849	13	4.6	2256	2	US-08-971-217-2	Sequence 2, Appl	C 922	13	4.6	2688	5	US-08-441-751-3	Sequence 3, Appl
C 850	13	4.6	2256	2	US-09-350-600-2	Sequence 2, Appl	C 923	13	4.6	2688	5	PCT-US92-02521-3	Sequence 3, Appl
C 851	13	4.6	2265	2	US-08-940-332-1	Sequence 1, Appl	C 924	13	4.6	2693	4	US-09-141-212-5	Sequence 5, Appl
C 852	13	4.6	2281	1	US-08-164-614A-1	Sequence 1, Appl	C 925	13	4.6	2693	4	US-09-561-138-5	Sequence 5, Appl
C 853	13	4.6	2281	2	US-08-456-489B-1	Sequence 1, Appl	C 926	13	4.6	2709	4	US-09-251-372-1	Sequence 1, Appl
C 854	13	4.6	2281	5	PCT-US93-01720-1	Sequence 1, Appl	C 927	13	4.6	2710	4	US-09-232-200-44	Sequence 44, Appl
C 855	13	4.6	2295	2	US-08-842-842-6	Sequence 6, Appl	C 928	13	4.6	2710	4	US-09-232-200-70	Sequence 70, Appl
C 856	13	4.6	2301	4	US-09-232-191-8	Sequence 8, Appl	C 929	13	4.6	2710	4	US-09-232-197-44	Sequence 44, Appl
C 857	13	4.6	2301	4	US-09-232-200-8	Sequence 8, Appl	C 930	13	4.6	2710	4	US-09-232-197-70	Sequence 70, Appl
C 858	13	4.6	2301	3	US-09-233-197-8	Sequence 8, Appl	C 931	13	4.6	2726	1	US-08-461-823-1	Sequence 1, Appl
C 859	13	4.6	2305	3	US-08-526-136-1	Sequence 3, Appl	C 932	13	4.6	2726	1	US-09-188-930-213	Sequence 213, App
C 860	13	4.6	2311	3	US-08-528-136-3	Sequence 3, Appl	C 933	13	4.6	2729	2	US-08-874-186-42	Sequence 42, Appl
C 861	13	4.6	2319	3	US-09-058-489-90	Sequence 90, Appl	C 934	13	4.6	2808	3	US-08-870-126-7	Sequence 7, Appl
C 862	13	4.6	2341	3	US-09-187-049-11	Sequence 11, Appl	C 935	13	4.6	2812	1	US-08-920-812-16	Sequence 16, Appl
C 863	13	4.6	2344	1	US-07-695-472B-35	Sequence 35, Appl	C 936	13	4.6	2812	1	US-08-921-177-16	Sequence 16, Appl
C 864	13	4.6	2347	1	US-08-250-740-32	Sequence 32, Appl	C 937	13	4.6	2812	1	US-08-921-177-16	Sequence 16, Appl
C 865	13	4.6	2347	1	US-07-695-472B-34	Sequence 1, Appl	C 938	13	4.6	2812	1	US-08-362-877C-16	Sequence 16, Appl
C 866	13	4.6	2347	1	US-07-695-472B-36	Sequence 36, Appl	C 939	13	4.6	2812	2	US-08-920-828-16	Sequence 16, Appl
C 867	13	4.6	2347	1	US-07-695-472B-35	Sequence 35, Appl	C 940	13	4.6	2815	3	US-09-214-564A-1	Sequence 1, Appl
C 868	13	4.6	2369	2	US-08-883-534-4	Sequence 4, Appl	C 941	13	4.6	2823	2	US-08-398-008A-1	Sequence 1, Appl
C 869	13	4.6	2369	1	US-09-204-764-4	Sequence 1, Appl	C 942	13	4.6	2823	2	US-08-893-333-1	Sequence 1, Appl
C 870	13	4.6	2382	1	US-08-256-938-1	Sequence 1, Appl	C 943	13	4.6	2848	4	US-08-464-954A-2	Sequence 2, Appl
C 871	13	4.6	2395	4	US-08-446-935-7	Sequence 7, Appl	C 944	13	4.6	2855	1	US-07-923-976-5	Sequence 5, Appl
C 872	13	4.6	2399	4	US-09-385-801-1	Sequence 1, Appl	C 945	13	4.6	2880	2	US-08-500-857A-3	Sequence 3, Appl
C 873	13	4.6	2419	1	US-08-428-415-1	Sequence 1, Appl	C 946	13	4.6	2906	1	US-08-554-612C-49	Sequence 49, Appl
C 874	13	4.6	2419	1	US-08-379-685-1	Sequence 1, Appl	C 947	13	4.6	2928	2	US-08-462-481-1	Sequence 1, Appl
C 875	13	4.6	2419	2	US-08-854-029-1	Sequence 1, Appl	C 948	13	4.6	2928	2	US-08-436-771-1	Sequence 1, Appl
C 876	13	4.6	2419	4	US-08-428-762-1	Sequence 1, Appl	C 949	13	4.6	2928	2	US-08-436-771-3	Sequence 3, Appl
C 877	13	4.6	2420	1	US-08-073-383-1	Sequence 1, Appl	C 950	13	4.6	2928	2	US-08-434-998-1	Sequence 1, Appl
C 878	13	4.6	2420	3	US-08-328-239A-1	Sequence 1, Appl	C 951	13	4.6	2928	2	US-08-434-998-3	Sequence 3, Appl
C 879	13	4.6	2420	5	PCT-US93-00227-1	Sequence 1, Appl	C 952	13	4.6	2928	2	US-08-487-797-1	Sequence 1, Appl
C 880	13	4.6	2420	5	PCT-US94-06365-1	Sequence 1, Appl	C 953	13	4.6	2928	2	US-08-487-797-3	Sequence 3, Appl
C 881	13	4.6	2420	5	PCT-US95-13661-1	Sequence 1, Appl	C 954	13	4.6	2928	2	US-08-701-005A-1	Sequence 1, Appl
C 882	13	4.6	2455	2	US-08-256-938-3	Sequence 3, Appl	C 955	13	4.6	2928	2	US-08-479-995-1	Sequence 1, Appl
C 883	13	4.6	2455	2	US-08-797-689-15	Sequence 15, Appl	C 956	13	4.6	2928	5	PCT-US95-02058-1	Sequence 1, Appl
C 884	13	4.6	2458	3	US-09-071-101-5	Sequence 5, Appl	C 957	13	4.6	2928	5	PCT-US95-02058-3	Sequence 3, Appl
C 885	13	4.6	2458	3	US-09-369-618-6	Sequence 6, Appl	C 958	13	4.6	2931	3	US-08-943-956A-1	Sequence 1, Appl
C 886	13	4.6	2458	3	US-09-369-617-6	Sequence 6, Appl	C 959	13	4.6	2943	1	US-07-923-976-3	Sequence 3, Appl
C 887	13	4.6	2463	1	US-08-370-542-1	Sequence 1, Appl	C 960	13	4.6	2952	4	US-08-306-691B-54	Sequence 54, Appl
C 888	13	4.6	2463	1	US-08-542-358-1	Sequence 1, Appl	C 961	13	4.6	3006	4	US-09-552-351-1	Sequence 1, Appl
C 889	13	4.6	2463	3	US-09-018-351-1	Sequence 1, Appl	C 962	13	4.6	3024	1	US-07-923-976-7	Sequence 7, Appl
C 890	13	4.6	2484	4	US-09-276-531-46	Sequence 46, Appl	C 963	13	4.6	3040	4	US-09-378-255-3	Sequence 3, Appl
C 891	13	4.6	2556	1	US-08-725-012-1	Sequence 1, Appl	C 964	13	4.6	3040	4	US-09-715-336-3	Sequence 3, Appl
C 892	13	4.6	2581	2	US-09-013-634-1	Sequence 1, Appl	C 965	13	4.6	3059	1	US-08-162-809-3	Sequence 3, Appl
C 893	13	4.6	2582	2	US-08-480-994-6	Sequence 6, Appl	C 966	13	4.6	3097	4	US-09-282-147-38	Sequence 38, Appl
C 894	13	4.6	2582	2	US-08-616-844-6	Sequence 6, Appl	C 967	13	4.6	3103	3	US-08-826-246-3	Sequence 3, Appl
C 895	13	4.6	2582	2	US-08-599-654-6	Sequence 6, Appl	C 968	13	4.6	3103	3	US-08-944-495-3	Sequence 3, Appl
C 896	13	4.6	2582	2	US-08-485-573-6	Sequence 6, Appl	C 969	13	4.6	3103	3	US-09-126-640-2	Sequence 2, Appl
C 897	13	4.6	2582	3	US-08-944-868A-6	Sequence 6, Appl	C 970	13	4.6	3103	4	US-08-925-588-3	Sequence 3, Appl
C 898	13	4.6	2582	3	US-08-944-423A-6	Sequence 6, Appl	C 971	13	4.6	3111	3	US-09-487-444-3	Sequence 3, Appl
C 899	13	4.6	2582	3	US-08-925-743-6	Sequence 6, Appl	C 972	13	4.6	3117	1	US-08-172-331B-3	Sequence 3, Appl
C 900	13	4.6	2582	3	US-08-944-496-6	Sequence 6, Appl	C 973	13	4.6	3125	1	US-08-162-809-19	Sequence 19, Appl
C 901	13	4.6	2582	4	US-08-925-767-6	Sequence 6, Appl	C 974	13	4.6	3236	4	US-08-927-219-7	Sequence 7, Appl
C 902	13	4.6	2588	3	US-08-796-414B-6	Sequence 6, Appl	C 975	13	4.6	3238	4	US-08-927-219-1	Sequence 1, Appl
C 903	13	4.6	2594	3	US-08-989-385-2	Sequence 2, Appl	C 976	13	4.6	3238	4	US-08-927-219-3	Sequence 3, Appl

977 13 4.6 3239 4 US-08-927-219-5 Sequence 5, Appli
978 13 4.6 3244 2 US-08-321-670-1 Sequence 1, Appli
979 13 4.6 3255 2 US-08-916-917-11 Sequence 11, Appl
980 13 4.6 3255 3 US-09-225-170-11 Sequence 11, Appl
981 13 4.6 3306 1 US-08-261-206A-71 Sequence 71, Appl
982 13 4.6 3321 1 US-08-486-270-1 Sequence 1, Appli
983 13 4.6 3321 3 US-08-367-264-1 Sequence 1, Appli
984 13 4.6 3330 1 US-08-072-574-1 Sequence 10, Appl
985 13 4.6 3429 1 US-08-097-997A-10 Sequence 10, Appl
986 13 4.6 3429 1 US-08-665-574C-10 Sequence 10, Appl
987 13 4.6 3429 4 US-08-946-994-10 Sequence 10, Appl
988 13 4.6 3447 2 US-08-313-185-57 Sequence 57, Appl
989 13 4.6 3447 3 US-09-082-614A-57 Sequence 57, Appl
990 13 4.6 3453 1 US-07-813-593-1 Sequence 1, Appli
991 13 4.6 3453 1 US-07-977-451-1 Sequence 1, Appli
992 13 4.6 3453 1 US-07-946-507-1 Sequence 1, Appli
993 13 4.6 3453 1 US-08-252-517-1 Sequence 1, Appli
994 13 4.6 3453 1 US-07-906-397A-1 Sequence 1, Appli
995 13 4.6 3453 1 US-08-601-891-1 Sequence 1, Appli
996 13 4.6 3453 2 US-09-021-324-1 Sequence 1, Appli
997 13 4.6 3453 5 PCT-US92-02750-1 Sequence 1, Appli
998 13 4.6 3453 5 PCT-US92-05401-1 Sequence 1, Appli
999 13 4.6 3453 5 PCT-US92-09893-1 Sequence 1, Appli
1000 13 4.6 3456 3 US-08-704-711A-8 Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-808-982-3
; Sequence 3, Application US/08808982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808, 982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2831 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-808-982-3

Query Match 6.4%; Score 18; DB 2; Length 2831;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 ctggagctgctgctgctg 36
|||||
Db 1469 CTGGGCTGCTGCTGCTG 1486

RESULT 2

US-09-306-902A-3
; Sequence 3, Application US/09306902A
; Patent No. 6277585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/306, 902A
; FILING DATE: 07-May-1999
; CLASSIFICATION: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2831 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-306-902A-3

Query Match 6.4%; Score 18; DB 4; Length 2831;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 ctggagctgctgctgctg 36
|||||
Db 1469 CTGGGCTGCTGCTGCTG 1486

RESULT 3

US-08-036-555B-44/C
; Sequence 44, Application US/08036555B
; Patent No. 5530109
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew; Stroobant, Paul;
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;
; APPLICANT: Chen, Maio Su; Hiles, Ian

```

: TITLE OF INVENTION: Glial Mitogenic Factors, Their
: TITLE OF INVENTION: Preparation and Use
: NUMBER OF SEQUENCES: 184
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Felife & Lynch
: STREET: 805 Third Avenue
: CITY: New York City
: STATE: New York
: COUNTRY: USA
: ZIP: 10022
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
: COMPUTER: IBM
: OPERATING SYSTEM: PC-DOS
: SOFTWARE: Wordperfect
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/036,555B
: FILING DATE: 24-MAR-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/965,173
: FILING DATE: 23-OCT-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/940,389
: FILING DATE: 03-SEP-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/907,138
: FILING DATE: 30-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/863,703
: FILING DATE: 03-APRIL-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: U.K. 91 07566.3
: FILING DATE: 10-APRIL-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Tsai, Christine H.
: REGISTRATION NUMBER: 34,266
: REFERENCE/DOCKET NUMBER: LUD 5250.4
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 688-9200
: TELEFAX: (212) 838-3884
: INFORMATION FOR SEQ. ID NO: 44:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 569
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: Linear
:
: US-08-036-555B-44
:
: Query Match 5.7%: Score 16; DB 1; Length 569;
: Best Local Similarity 100.0%; Pred. No. 32;
: Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: Oy 66 tgglttgcagtagc 81
: |||||||||||||||
: Db 103 TGGTTTGCAGTAGC 88
:
: RESULT 4
: US-08-469-569-44/c
: Sequence 44, Application US/08469569
: Patent No. 5606032
: GENERAL INFORMATION:
: APPLICANT: Goodearl, Andrew; Stroobant, Paul;
: APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
: APPLICANT: Chen, Maio Su; Hiles, Ian
: TITLE OF INVENTION: Glial Mitogenic Factors, Their
: TITLE OF INVENTION: Preparation and Use
: NUMBER OF SEQUENCES: 184
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Felife & Lynch
: STREET: 805 Third Avenue
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10022

```

```

: CITY: New York City
: STATE: New York
: COUNTRY: USA
: ZIP: 10022
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
: COMPUTER: IBM
: OPERATING SYSTEM: PC-DOS
: SOFTWARE: Wordperfect
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/469,569
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/036,555
: FILING DATE: 24-MAR-1993
: APPLICATION NUMBER: 07/965,173
: FILING DATE: 23-OCT-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/940,389
: FILING DATE: 03-SEP-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/907,138
: FILING DATE: 30-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/863,703
: FILING DATE: 03-APRIL-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: U.K. 91 07566.3
: FILING DATE: 10-APRIL-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Tsai, Christine H.
: REGISTRATION NUMBER: 34,266
: REFERENCE/DOCKET NUMBER: LUD 5250.4
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 688-9200
: TELEFAX: (212) 838-3884
: INFORMATION FOR SEQ. ID NO: 44:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 569
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: Linear
:
: US-08-469-569-44
:
: Query Match 5.7%: Score 16; DB 1; Length 569;
: Best Local Similarity 100.0%; Pred. No. 32;
: Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: Oy 66 tgglttgcagtagc 81
: |||||||||||||||
: Db 103 TGGTTTGCAGTAGC 88
:
: RESULT 5
: US-08-249-322A-44/c
: Sequence 44, Application US/08249322A
: Patent No. 5716930
: GENERAL INFORMATION:
: APPLICANT: Goodearl, Andrew; Stroobant, Paul;
: APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
: APPLICANT: Chen, Maio Su; Hiles, Ian
: TITLE OF INVENTION: Glial Mitogenic Factors, Their
: TITLE OF INVENTION: Preparation and Use
: NUMBER OF SEQUENCES: 184
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Felife & Lynch
: STREET: 805 Third Avenue
: CITY: New York City
: STATE: New York
: COUNTRY: USA
: ZIP: 10022

```

```

:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
: COMPUTER: IBM
: OPERATING SYSTEM: PC-DOS
: SOFTWARE: Wordperfect
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/249,322A
: FILING DATE: 26-MAY-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/036,555
: FILING DATE: 24-MAR-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/965,173
: FILING DATE: 23-OCT-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/940,389
: FILING DATE: 03-SEP-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/907,138
: FILING DATE: 30-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/863,703
: FILING DATE: 03-APRIL-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: U.K. 91 07566.3
: FILING DATE: 10-APRIL-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Tsai, Christine H.
: REGISTRATION NUMBER: 34,266
: REFERENCE/DOCKET NUMBER: LUD 250.4
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 688-9200
: TELEFAX: (212) 838-3884
: INFORMATION FOR SEQ ID NO: 44:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 569
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: US-08-249-322A-44
:
: Query Match 5.7%; Score 16; DB 1; Length 569;
: Best Local Similarity 100.0%; Pred. No. 32;
: Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: QY 66 tggtttcagtaggc 81
: |||||||||||||||
: Db 103 TCGTTTTCAGTAGGC 88
:
: RESULT 6
: US-08-469-526A-44/C
: Sequence 44, Application US/08469526A
: Patent No. 5792849
: GENERAL INFORMATION:
: APPLICANT: Goodearl, Andrew
: APPLICANT: Stroobant, Paul
: APPLICANT: Minghetti, Luisa
: APPLICANT: Waterfield, Michael
: APPLICANT: Marchionni, Mark
: APPLICANT: Chen, Maio Su
: APPLICANT: Hiles, Ian
: TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
: PREPARATION AND USE
: NUMBER OF SEQUENCES: 187
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Clark & Ebling LLP
: STREET: 176 Federal Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA

```

```

:
: ZIP: 02110
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/469,526A
: FILING DATE: 06 June 1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/036,555
: FILING DATE: 24-MAR-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/965,173
: FILING DATE: 23-OCT-1992
: APPLICATION NUMBER: 07/940,389
: FILING DATE: 03-SEP-1992
: APPLICATION NUMBER: 07/907,138
: FILING DATE: 03-JUN-1992
: APPLICATION NUMBER: 07/863,703
: FILING DATE: 03-APRIL-1992
: APPLICATION NUMBER: U.K. 91 07566.3
: FILING DATE: 10-APR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Bieker-Brady, Kristina
: REGISTRATION NUMBER: 39,109
: REFERENCE/DOCKET NUMBER: 04585/00200A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-428-0200
: TELEFAX: 617-428-7045
: INFORMATION FOR SEQ ID NO: 44:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 569
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: US-08-469-526A-44
:
: Query Match 5.7%; Score 16; DB 1; Length 569;
: Best Local Similarity 100.0%; Pred. No. 32;
: Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: QY 66 tggtttcagtaggc 81
: |||||||||||||||
: Db 103 TCGTTTTCAGTAGGC 88
:
: RESULT 7
: US-08-734-591A-44/C
: Sequence 44, Application US/08734591A
: Patent No. 5854220
: GENERAL INFORMATION:
: APPLICANT: Goodearl, Andrew
: APPLICANT: Stroobant, Paul
: APPLICANT: Minghetti, Luisa
: APPLICANT: Waterfield, Michael
: APPLICANT: Hiles, Ian
: APPLICANT: Marchionni, Mark
: APPLICANT: Chen, Maio
: TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
: PREPARATION AND USE
: NUMBER OF SEQUENCES: 187
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Clark & Ebling LLP
: STREET: 176 Federal Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02110
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
: COMPUTER: IBM Compatible Pentium

```



```
OPERATING SYSTEM: Windows95
SOFTWARE: Wordperfect (Version 7.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,591A
FILING DATE: 22-OCT-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/470,335
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/036,555
FILING DATE: 03-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,173
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/907,138
FILING DATE: 30-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 91 07566.3
FILING DATE: 10-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bieker-Brady, Kristina
REGISTRATION NUMBER: 39,109
REFERENCE/DOCKET NUMBER: 04585/00200P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 428-0200
TELEFAX: (617) 428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 569
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-734-591A-44

Query Match          5.7%; Score 16; DB 2; Length 569;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 tggtttcagtaggc 81
    |||
Db 103 TGGTTTGCAGTAGGC 88

RESULT 8
US-08-469-660-181/C
Sequence 181, Application US/08469660
Patent No. 5876973
GENERAL INFORMATION:
APPLICANT: Gwynne, David L.; Marchionni, Mark;
APPLICANT: McBurney, Robert N.
TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
ZIP: 0211-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
```

```
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,660
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/011,396
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/984,085
FILING DATE: 01-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/951,747
FILING DATE: 25-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/927,337
FILING DATE: 10-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 04585/017004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: 200154
INFORMATION FOR SEQ ID NO: 181:
SEQUENCE CHARACTERISTICS:
LENGTH: 569
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-469-660-181

Query Match          5.7%; Score 16; DB 2; Length 569;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 tggtttcagtaggc 81
    |||
Db 103 TGGTTTGCAGTAGGC 88

RESULT 9
US-08-341-018-43/C
Sequence 43, Application US/08341018A
Patent No. 6087323
GENERAL INFORMATION:
APPLICANT: Gwynne, David L.
APPLICANT: Mahanthappa, Nagesh K.
APPLICANT: Marchionni, Mark A.
APPLICANT: Birmingham-McDonogh, Olivia
APPLICANT: Goldin, Stanley M.
APPLICANT: McBurney, Robert N.
TITLE OF INVENTION: USE OF NEUREGULINS AS MODULATORS OF
FILE REFERENCE: 04585/041001
CURRENT APPLICATION NUMBER: US/08/341,018A
CURRENT FILING DATE: 1994-11-17
NUMBER OF SEQ ID NOS: 87
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43
LENGTH: 569
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(569)
US-08-341-018-43

Query Match          5.7%; Score 16; DB 3; Length 569;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 66 tggcttcgacgtagc 81
|||||
DB 103 TGGTTTGCAGTAGGC 88

RESULT 10
US-08-470-335-44/C
; Sequence 44, Application US/08470335F
; Patent No. 6147190
; GENERAL INFORMATION:
; APPLICANT: GOODEARL, ANDREW
; APPLICANT: STROOBANT, PAUL
; APPLICANT: MINGHETTI, LUISA
; APPLICANT: WATERFIELD, MICHAEL
; APPLICANT: MARCHIONNI, MARK
; APPLICANT: CHEN, MARIO S.
; APPLICANT: HILES, IAN
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; TITLE OF INVENTION: PREPARATION AND USE
; FILE REFERENCE: 04585/00200B
; CURRENT APPLICATION NUMBER: US/08/470,335F
; EARLIER FILING DATE: 1995-06-06
; EARLIER APPLICATION NUMBER: 08/036,555
; EARLIER FILING DATE: 1993-03-24
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(569)
US-08-470-335-44

Query Match 5.7%; Score 16; DB 3; Length 569;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 tggcttcgacgtagc 81
|||||
DB 103 TGGTTTGCAGTAGGC 88

RESULT 11
US-08-735-021-44/C
; Sequence 44, Application US/08735021B
; Patent No. 6194377
; GENERAL INFORMATION:
; APPLICANT: GOODEARL, ANDREW
; APPLICANT: STROOBANT, PAUL
; APPLICANT: MINGHETTI, LUISA
; APPLICANT: WATERFIELD, MICHAEL
; APPLICANT: MARCHIONNI, MARK
; APPLICANT: CHEN, MARIO S.
; APPLICANT: HILES, IAN
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; TITLE OF INVENTION: PREPARATION AND USE
; FILE REFERENCE: 04585/00200L
; CURRENT APPLICATION NUMBER: US/08/735,021B
; EARLIER FILING DATE: 1996-10-22
; EARLIER APPLICATION NUMBER: 08/472,065
; EARLIER FILING DATE: 1995-06-06
; EARLIER APPLICATION NUMBER: 08/036,555
; EARLIER FILING DATE: 1993-03-24
; EARLIER APPLICATION NUMBER: 07/965,173
; EARLIER FILING DATE: 1992-10-23
; EARLIER APPLICATION NUMBER: 07/940,389
; EARLIER FILING DATE: 1992-09-03
; EARLIER APPLICATION NUMBER: 07/907,138
; EARLIER FILING DATE: 1992-06-30

; EARLIER APPLICATION NUMBER: 07/863,703
; EARLIER FILING DATE: 1992-04-03
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-735-021-44

Query Match 5.7%; Score 16; DB 4; Length 569;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 tggcttcgacgtagc 81
|||||
DB 103 TGGTTTGCAGTAGGC 88

RESULT 12
US-08-734-664A-44/C
; Sequence 44, Application US/08734664A
; Patent No. 6204241
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew
; APPLICANT: Stroobant, Paul
; APPLICANT: Minghetti, Luisa
; APPLICANT: Waterfield, Michael
; APPLICANT: Marchionni, Mark
; APPLICANT: Chen, Mario
; APPLICANT: Hiles, Ian
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; TITLE OF INVENTION: PREPARATION AND USE
; NUMBER OF SEQUENCES: 187
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible Pentium
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,664A
; FILING DATE: 22-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/249,322
; FILING DATE: 26-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 91 07566.3
; FILING DATE: 10-APR-1991
; ATTORNEY/AGENT INFORMATION:

```
NAME: Bieker-Brady, Kristina
REGISTRATION NUMBER: 39,109
REFERENCE/DOCKET NUMBER: 04585/00200J
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 428-0200
TELEFAX: (617) 428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 569
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-734-664A-44
```

```
Query Match          5.7%: Score 16; DB 4; Length 569;
Best Local Similarity 100.0%: Pred. No. 32;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 66 tggatttcagtaggc 81
    |||
Db 103 TGGTTTGCAGTAGGC 88
```

```
RESULT 13
US-08-470-339-44/C
Sequence 44, Application US/08470339C
Patent No. 6232286
GENERAL INFORMATION:
APPLICANT: GOODEARL, ANDREW
APPLICANT: STROOBANT, PAUL
APPLICANT: MINGHETTI, LUISA
APPLICANT: WATERFIELD, MICHAEL
APPLICANT: MARCHIONNI, MARK
APPLICANT: CHEN, MARIO S.
APPLICANT: HILES, IAN
TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
FILE REFERENCE: 04585/002008
CURRENT APPLICATION NUMBER: US/08/470,339C
CURRENT FILING DATE: 1995-06-06
EARLIER APPLICATION NUMBER: 08/036,555
EARLIER FILING DATE: 1993-03-24
EARLIER APPLICATION NUMBER: 07/940,389
EARLIER FILING DATE: 1992-09-03
EARLIER APPLICATION NUMBER: 07/907,138
EARLIER FILING DATE: 1992-06-30
EARLIER APPLICATION NUMBER: 07/863,703
EARLIER FILING DATE: 1992-04-03
EARLIER APPLICATION NUMBER: 91 07566.3 GB
EARLIER FILING DATE: 1989-04-10
NUMBER OF SEQ ID NOS: 226
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 44
LENGTH: 569
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(569)
US-08-470-339-44
```

```
Query Match          5.7%: Score 16; DB 4; Length 569;
Best Local Similarity 100.0%: Pred. No. 32;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 66 tggatttcagtaggc 81
    |||
Db 103 TGGTTTGCAGTAGGC 88
```

```
RESULT 14
PCT-US94-05083C-177/C
Sequence 177, Application PC/TUS9405083C
GENERAL INFORMATION:
APPLICANT: Robert Sklar, Mark Marchionni,
APPLICANT: David I. Gwynne
TITLE OF INVENTION: METHODS FOR ALTERING
NUMBER OF SEQUENCES: 185
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360
MEDIUM TYPE: kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Mordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05083C
FILING DATE: 06-MAY-94
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/209,204
FILING DATE: 08-MAR-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/059,022
FILING DATE: 06-MAY-93
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 04585/028W01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SRO ID NO: 177:
SEQUENCE CHARACTERISTICS:
LENGTH: 569
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US94-05083C-177
```

```
Query Match          5.7%: Score 16; DB 5; Length 569;
Best Local Similarity 100.0%: Pred. No. 32;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 66 tggatttcagtaggc 81
    |||
Db 103 TGGTTTGCAGTAGGC 88
```

```
RESULT 15
PCT-US95-06846A-44/C
Sequence 44, Application PC/TUS9506846A
GENERAL INFORMATION:
APPLICANT: Goodearl, Andrew David; Stroobant, Paul;
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;
APPLICANT: Chen, Mario Su; Hiles, Ian
TITLE OF INVENTION: Glial Mitogenic Factors, Their
NUMBER OF SEQUENCES: 178
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
```

```

1 ZIP: 10022
2 COMPUTER READABLE FORM:
3 MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
4 COMPUTER: IBM
5 OPERATING SYSTEM: PC-DOS
6 SOFTWARE: Wordperfect
7 CURRENT APPLICATION DATA:
8 APPLICATION NUMBER: PCT/US95/06846A
9 FILING DATE: 25-MAY-1995
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER: 08/249,322
12 FILING DATE: 26-MAY-1994
13 CLASSIFICATION:
14 PRIOR APPLICATION DATA:
15 APPLICATION NUMBER: 08/036,555
16 FILING DATE: 24-MAR-1993
17 PRIOR APPLICATION DATA:
18 APPLICATION NUMBER: 07/965,173
19 FILING DATE: 23-OCT-1992
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: 07/940,389
22 FILING DATE: 03-SEP-1992
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: 07/907,138
25 FILING DATE: 30-JUN-1992
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: 07/863,703
28 FILING DATE: 03-APRIL-1992
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: U.K. 91 07566.3
31 FILING DATE: 10-APRIL-1991
32 ATTORNEY/AGENT INFORMATION:
33 NAME: Hanson, Norman D.
34 REGISTRATION NUMBER: 30,946
35 REFERENCE/DOCKET NUMBER: LUD 5250.5
36 TELECOMMUNICATION INFORMATION:
37 TELEPHONE: (212) 688-9200
38 TELEFAX: (212) 838-3884
39 INFORMATION FOR SEQ ID NO: 44:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 569
42 TYPE: nucleic acid
43 STRANDEDNESS: single
44 TOPOLOGY: linear
45
46 PCT-US95-06846A-44

```

Query Match	5.7%	Score 16	DB 5	Length 569
Best Local Similarity	100.0%	Pred. No. 32		
Matches	16	Conservative	0	Mismatches 0
				Indels 0
				Gaps 0
Qy	66	tggttttcgaatgac	81	
Db	103	tcgtrtttcagatggcc	88	

Search completed: February 13, 2002, 20:05:45
Job time: 18172 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 18:47:39 : Search time 2841.05 Seconds
(without alignments)
1066.616 Million cell updates/sec

Title: US-09-516-493-11

Perfect score: 282

Sequence: 1 gagcctgcgtatgttcaacct.....tccctcagcgtccttccacg 282

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 11351937 seqs, 5372889281 residues

Word size: 0

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database: EST:*

1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: qb_est1:*
11: qb_est2:*
12: qb_hlc:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	180	63.8	319	11 H34451	H34451 EST111406 R
2	150	17.7	595	11 BG079217	BG079217 H3037C03-
3	50	17.7	823	11 BF120840	BF120840 601757442
4	48	17.0	917	11 BF140667	BF140667 601786917
5	44	15.6	440	11 BF774206	BF774206 283798 MA
6	41	14.5	327	11 BF742266	BF742266 RCI-BRN04
7	41	14.5	358	10 AA627408	AA627408 nq49Q04.s
8	41	14.5	496	10 AA248655	AA248655 2820759.3
9	41	14.5	712	11 BG290178	BG290178 602385265
10	41	14.5	934	11 BE910478	BE910478 601501005
11	41	14.5	949	11 BG749509	BG749509 602707648
12	30	10.6	480	10 AW249090	AW249090 2820759.5

13	29	10.3	538	11 BG672321	BG672321 DRNCFD10
14	27	9.6	546	10 BE231636	BE231636 136437 MA
15	26	9.2	752	10 AL565390	AL565390 AL565390
16	26	9.2	796	13 AZ985752	AZ985752 2M0267011
17	24	8.5	235	11 BF146289	BF146289 EST00181
18	23	8.2	279	11 H46443	H46443 Y016D08.r1
19	23	8.2	830	11 BG700749	BG700749 602681616
20	23	8.2	974	10 AL534131	AL534131 AL534131
21	23	8.2	991	11 B1334832	B1334832 602998959
22	21	7.4	694	13 AO689602	AO689602 nbx0080C
23	21	7.4	335	13 AO027940	AO027940 CTF-HSP-2
24	20	7.1	355	10 AA526198	AA526198 n159D01.s
25	20	7.1	373	13 AO034369	AO034369 CTF-HSP-2
26	20	7.1	394	11 H07947	H07947 y186a04.r1
27	20	7.1	487	11 BF333736	BF333736 OYO-CS001
28	20	7.1	710	13 AO027042	AO027042 CTF-HSP-2
29	20	7.1	911	11 BF311341	BF311341 601896741
30	19	6.7	262	10 AA379176	AA379176 EST91963
31	19	6.7	367	10 AU179337	AU179337 PM1-HT045
32	19	6.7	377	10 BF351127	BF351127 PM1-HT045
33	19	6.7	377	10 AW513410	AW513410 XO4-C08.x
34	19	6.7	380	11 BE991517	BE991517 U1-M-B21-
35	19	6.7	395	10 BE707361	BE707361 OY2-HT049
36	19	6.7	423	10 BE482284	BE482284 168018 BA
37	19	6.7	435	13 AO007897	AO007897 CTF-HSP-2
38	19	6.7	441	10 BE777958	BE777958 601463445
39	19	6.7	457	10 AA405891	AA405891 zu57d11.r
40	19	6.7	459	10 BE482283	BE482283 168017 BA
41	19	6.7	466	10 A1364016	A1364016 gw35e11.x
42	19	6.7	473	10 AW043663	AW043663 w774g11.x
43	19	6.7	493	10 AA479596	AA479596 z142901.r
44	19	6.7	500	10 A1274803	A1274803 gw01h05.x
45	19	6.7	508	10 AW266681	AW266681 L48-11T3
46	19	6.7	508	13 BH060704	BH060704 RPCI-24-3
47	19	6.7	510	10 AW949842	AW949842 EST631912
48	19	6.7	513	10 AL600774	AL600774 DFEZP313C
49	19	6.7	514	11 BF819856	BF819856 MRI-RT002
50	19	6.7	522	10 BE208202	BE208202 bb62d12.x
51	19	6.7	528	13 AO709148	AO709148 HS-5329_B
52	19	6.7	536	10 AW513427	AW513427 XO4-F06.x
53	19	6.7	538	10 BE208410	BE208410 bb24h11.x
54	19	6.7	546	10 AU158997	AU158997 AU158997
55	19	6.7	566	10 AW129662	AW129662 x622D09.x
56	19	6.7	567	10 A1889449	A1889449 w04e12.x
57	19	6.7	572	10 AW572480	AW572480 xu65d10.x
58	19	6.7	578	10 AW779980	AW779980 hn91e11.x
59	19	6.7	581	10 A1816837	A1816837 w334Q02.x
60	19	6.7	581	13 AZ605516	AZ605516 1M0426F21
61	19	6.7	593	11 BF351120	BF351120 PM1-HT045
62	19	6.7	595	10 AU180129	AU180129 AU180129
63	19	6.7	595	10 BE311406	BE311406 601142761
64	19	6.7	596	10 AU180108	AU180108 AU180108
65	19	6.7	613	10 BE037163	BE037163 MP16R02 M
66	19	6.7	614	10 A1690736	A1690736 tX88901.x
67	19	6.7	615	10 A1561239	A1561239 lq27f01.x
68	19	6.7	616	10 AW053298	AW053298 L30-1390T
69	19	6.7	621	11 BF582949	BF582949 602097805
70	19	6.7	622	10 AA653391	AA653391 ad65h11.s
71	19	6.7	627	10 BE037164	BE037164 MP16R03 M
72	19	6.7	627	11 B1000856	B1000856 MK3-HN006
73	19	6.7	650	10 AW053562	AW053562 L30-2112T
74	19	6.7	672	13 AZ122022	AZ122022 RPCI-23-1
75	19	6.7	690	11 BF480389	BF480389 LO-2321T3
76	19	6.7	702	10 AA994967	AA994967 ou89h11.s
77	19	6.7	710	10 AU170414	AU170414 AU170414
78	19	6.7	715	10 AU169291	AU169291 AU169291
79	19	6.7	716	10 AU172151	AU172151 AU172151
80	19	6.7	717	10 A1978894	A1978894 wT61d01.x
81	19	6.7	718	11 B1117844	B1117844 602866867
82	19	6.7	748	10 BE744040	BE744040 601576945
83	19	6.7	781	10 BE034212	BE034212 MH01C05 M
84	19	6.7	808	10 BE568142	BE568142 601341669
85	19	6.7	870	11 BF031238	BF031238 601560162

86	19	6.7	951	10	BE612638	601452477	C	159	17	6.0	309	11	BI286939	UI-R-CTOS
87	19	6.7	960	10	BE748835	601572070	C	160	17	6.0	310	10	AV387568	AV387568
C 88	19	6.7	1035	10	BE036587	MP01G03 M	C	161	17	6.0	311	10	BA434063	BA434063
C 89	19	6.7	1099	11	BC164921	602343546	C	162	17	6.0	312	17	AI180588	AI180588
C 90	18	6.4	114	90	BE477069	160710 BA	C	163	17	6.0	317	13	AZ481196	AZ481196
C 91	18	6.4	224	13	AO634020	RPCT-11-4	C	164	17	6.0	320	11	BC983254	BC983254
C 92	18	6.4	239	10	BE477070	160711 BA	C	165	17	6.0	321	11	BE544322	BE544322
C 93	18	6.4	250	10	AM508496	s131d03.y	C	166	17	6.0	326	13	B53504	B53504
C 94	18	6.4	261	11	BI131069	dafe86902.	C	167	17	6.0	331	11	AM867636	AM867636
C 95	18	6.4	265	10	AA701213	zj76h09.s	C	168	17	6.0	334	11	BC983262	BC983262
C 96	18	6.4	271	10	AV172244	AV172244	C	169	17	6.0	336	13	B46300	B46300
C 97	18	6.4	295	13	AO079652	CIT-HSP-2	C	170	17	6.0	340	10	AA131274	AA131274
C 98	18	6.4	301	11	BI004789	MR4-HN005	C	171	17	6.0	344	10	BE670840	BE670840
C 99	18	6.4	321	10	BE477831	161771 BA	C	172	17	6.0	370	11	N79473	N79473
C 100	18	6.4	363	11	BF441289	257508 MA	C	173	17	6.0	371	13	AO191188	AO191188
C 101	18	6.4	371	13	AO120937	HS-3073-B	C	174	17	6.0	384	13	AO493294	AO493294
C 102	18	6.4	412	11	BE805710	ss48906.y	C	175	17	6.0	385	11	BF715796	BF715796
C 103	18	6.4	428	10	AA112607	zn70d04.r	C	176	17	6.0	392	10	AA954830	AA954830
C 104	18	6.4	467	13	AO952651	Sheared D	C	177	17	6.0	397	10	AV633086	AV633086
C 105	18	6.4	468	10	AI481382	vg10a04.x	C	178	17	6.0	399	11	BG405718	BG405718
C 106	18	6.4	495	13	AO686339	nbxb0068E	C	179	17	6.0	399	13	B51398	B51398
C 107	18	6.4	502	13	AO446280	nbxb0069E	C	180	17	6.0	401	10	AV638993	AV638993
C 108	18	6.4	513	11	BI045456	MR3-FN020	C	181	17	6.0	409	13	AO093245	AO093245
C 109	18	6.4	514	13	AO721170	HS-5557_A	C	182	17	6.0	413	10	AV633107	AV633107
C 110	18	6.4	517	10	BE476250	158891 BA	C	183	17	6.0	416	10	AA599685	AA599685
C 111	18	6.4	521	10	AM760092	sl58a03.y	C	184	17	6.0	418	11	BE848881	BE848881
C 112	18	6.4	521	13	AO179923	HS-3203-B	C	185	17	6.0	418	11	BF954127	BF954127
C 113	18	6.4	522	11	BG747686	602705157	C	186	17	6.0	422	10	AV633175	AV633175
C 114	18	6.4	522	11	BG881558	sae85904.	C	187	17	6.0	425	11	W63217	W63217
C 115	18	6.4	523	13	AO347960	RPCT11-11	C	188	17	6.0	426	10	AV387800	AV387800
C 116	18	6.4	543	13	TA50601P	AI455868 T. brucei	C	189	17	6.0	428	10	AV636159	AV636159
C 117	18	6.4	552	11	BF279269	GA_Eb003	C	190	17	6.0	429	10	AV625176	AV625176
C 118	18	6.4	564	13	AO163640	nbxb0007B	C	191	17	6.0	432	10	AV625626	AV625626
C 119	18	6.4	564	13	TA346D10	AI496240 T. brucei	C	192	17	6.0	434	10	AL510735	AL510735
C 120	18	6.4	585	13	AO312982	RPCT11-11	C	193	17	6.0	436	10	AV638370	AV638370
C 121	18	6.4	591	10	AM387111	MR0-ST002	C	194	17	6.0	437	10	AV637725	AV637725
C 122	18	6.4	605	13	AO527380	AO527380 RPCT-11-3	C	195	17	6.0	438	10	AV632036	AV632036
C 123	18	6.4	618	13	AO257432	nbxb0018A	C	196	17	6.0	438	10	AV636080	AV636080
C 124	18	6.4	623	11	AO951399	Sheared D	C	197	17	6.0	439	10	AV637389	AV637389
C 125	18	6.4	623	11	BG443439	GA_Ea002	C	198	17	6.0	439	10	AV637418	AV637418
C 126	18	6.4	624	10	AM734248	sk87f05.y	C	199	17	6.0	439	10	AM652100	AM652100
C 127	18	6.4	641	11	BG809828	mgct001x1	C	200	17	6.0	440	10	AV624235	AV624235
C 128	18	6.4	643	13	A2507144	AM0348F07	C	201	17	6.0	440	10	AV637025	AV637025
C 129	18	6.4	658	10	AV762618	AV762618	C	202	17	6.0	441	10	BE121911	BE121911
C 130	18	6.4	659	13	A2220829	Sheared D	C	203	17	6.0	442	10	AV387762	AV387762
C 131	18	6.4	672	13	AO082692	RPCT11-54	C	204	17	6.0	442	10	AV634610	AV634610
C 132	18	6.4	688	13	AO256393	nbxb0016A	C	205	17	6.0	442	10	BE423487	BE423487
C 133	18	6.4	690	11	BG597732	EST496410	C	206	17	6.0	443	10	AV636534	AV636534
C 134	18	6.4	694	10	BE374096	BE374096 601228138	C	207	17	6.0	443	10	AV636933	AV636933
C 135	18	6.4	751	13	CNS0417Y	TEtracoon	C	208	17	6.0	444	10	AV640092	AV640092
C 136	18	6.4	811	11	BF306231	601893084	C	209	17	6.0	444	10	AV639647	AV639647
C 137	18	6.4	833	11	BF104897	601822626	C	210	17	6.0	444	11	N48157	N48157
C 138	18	6.4	882	11	BG445840	GA_Ea002	C	211	17	6.0	445	10	AV640973	AV640973
C 139	18	6.4	897	11	BG472444	602514135	C	212	17	6.0	447	10	AV635715	AV635715
C 140	18	6.4	921	13	AO899775	HS-3203-B	C	213	17	6.0	447	11	C03825	C03825
C 141	18	6.4	928	11	BG751850	602731401	C	214	17	6.0	448	10	AV635473	AV635473
C 142	18	6.4	934	13	BG197384	RST16626	C	215	17	6.0	448	10	AV636395	AV636395
C 143	18	6.4	981	13	CNS04WMU	AL310575 TEtracoon	C	216	17	6.0	448	10	AV640545	AV640545
C 144	18	6.4	1054	11	BF574031	602131863	C	217	17	6.0	449	10	AV387863	AV387863
C 145	18	6.4	1058	11	BI114745	602861781	C	218	17	6.0	449	13	AO241905	AO241905
C 146	17	6.0	159	10	AA482827	n148G12.s	C	219	17	6.0	451	10	AI232690	AI232690
C 147	17	6.0	248	11	BF049777	NXCI_110	C	220	17	6.0	451	11	BE805107	BE805107
C 148	17	6.0	253	11	F33520	HSPD27000 H	C	221	17	6.0	453	10	AV635681	AV635681
C 149	17	6.0	259	11	BC984591	CNO-16	C	222	17	6.0	456	10	AV620238	AV620238
C 150	17	6.0	293	10	BB597596	BB597596	C	223	17	6.0	456	11	R78268	R78268
C 151	17	6.0	278	10	AM035444	EST281182	C	224	17	6.0	456	13	AO219405	AO219405
C 152	17	6.0	285	13	F33906	HSPD28299 H	C	225	17	6.0	458	10	AV639111	AV639111
C 153	17	6.0	290	10	AA0060187	CIT-HSP-2	C	226	17	6.0	458	10	AM707507	AM707507
C 154	17	6.0	292	10	BB261704	BB261704	C	227	17	6.0	459	10	AV619662	AV619662
C 155	17	6.0	292	10	AL449512	AL449512	C	228	17	6.0	459	10	AV636664	AV636664
C 156	17	6.0	294	10	AT626149	f06f04.y	C	229	17	6.0	459	10	AM758460	AM758460
C 157	17	6.0	295	11	BC303512	f156g08.x	C	230	17	6.0	460	10	AV633079	AV633079
C 158	17	6.0	297	10	BB173214	BB173214	C	231	17	6.0	460	10	AV633682	AV633682

C 232	17	6.0	460	10	AV643245	AV643245	AV643245	C 305	17	6.0	514	11	N53303	N53303_ylz02c11.r1
C 233	17	6.0	465	10	AA181767	AA181767	AA181767	306	17	6.0	515	10	AV3889552	AV3889552
234	17	6.0	465	10	AV624611	AV624611	AV624611	307	17	6.0	516	10	AV632196	AV632196
235	17	6.0	465	10	AW758262	AW758262	AW758262	C 308	17	6.0	516	10	BE755506	BE755506
236	17	6.0	465	10	BE122121	BE122121	BE122121	309	17	6.0	517	10	AV630375	AV630375
237	17	6.0	466	10	AV633487	AV633487	AV633487	310	17	6.0	517	10	AV643007	AV643007
238	17	6.0	466	10	AV639403	AV639403	AV639403	311	17	6.0	517	10	AV643197	AV643197
239	17	6.0	468	10	AV625198	AV625198	AV625198	312	17	6.0	518	10	AV632244	AV632244
240	17	6.0	468	10	AV625464	AV625464	AV625464	313	17	6.0	518	10	AV633990	AV633990
241	17	6.0	468	10	AV635343	AV635343	AV635343	314	17	6.0	518	10	AV634032	AV634032
242	17	6.0	468	11	BC509593	BC509593	BC509593	315	17	6.0	519	10	AV633932	AV633932
C 243	17	6.0	468	11	BF929269	BF929269	BF929269	316	17	6.0	520	10	AV632077	AV632077
244	17	6.0	470	10	BE129384	BE129384	BE129384	317	17	6.0	521	10	AV635518	AV635518
245	17	6.0	472	10	AV635035	AV635035	AV635035	318	17	6.0	521	10	AV637302	AV637302
246	17	6.0	473	10	AV639353	AV639353	AV639353	319	17	6.0	522	10	AV624025	AV624025
247	17	6.0	474	10	AV625798	AV625798	AV625798	320	17	6.0	522	10	AV634471	AV634471
248	17	6.0	474	10	AV632381	AV632381	AV632381	321	17	6.0	522	10	AW758226	AW758226
249	17	6.0	474	10	AV634748	AV634748	AV634748	322	17	6.0	523	10	AV635634	AV635634
250	17	6.0	474	10	BE227755	BE227755	BE227755	C 323	17	6.0	524	11	BF045074	BF045074
251	17	6.0	475	10	AV638471	AV638471	AV638471	324	17	6.0	526	11	BE203293	BE203293
252	17	6.0	476	10	AV636267	AV636267	AV636267	325	17	6.0	526	11	BE803240	BE803240
253	17	6.0	477	10	AV635340	AV635340	AV635340	326	17	6.0	527	10	AV621954	AV621954
254	17	6.0	477	10	AV638442	AV638442	AV638442	C 327	17	6.0	529	10	BE422978	BE422978
255	17	6.0	477	10	AV641506	AV641506	AV641506	328	17	6.0	530	10	AV392172	AV392172
256	17	6.0	479	10	AV643500	AV643500	AV643500	329	17	6.0	531	10	AV621625	AV621625
C 257	17	6.0	479	10	AA420061	AA420061	AA420061	330	17	6.0	531	10	AV632243	AV632243
258	17	6.0	480	10	AV634612	AV634612	AV634612	331	17	6.0	532	10	AV634381	AV634381
259	17	6.0	480	11	BC852250	BC852250	BC852250	332	17	6.0	532	10	AV635748	AV635748
260	17	6.0	481	10	AV639312	AV639312	AV639312	333	17	6.0	532	10	AW772960	AW772960
261	17	6.0	481	10	AV645195	AV645195	AV645195	334	17	6.0	534	10	AV396800	AV396800
C 262	17	6.0	482	11	BF483537	BF483537	BF483537	C 335	17	6.0	534	11	BG868938	BG868938
263	17	6.0	483	10	AV624126	AV624126	AV624126	336	17	6.0	535	10	AI297452	AI297452
264	17	6.0	484	10	AV626645	AV626645	AV626645	337	17	6.0	539	10	AV386917	AV386917
265	17	6.0	484	10	AV631017	AV631017	AV631017	C 338	17	6.0	541	13	AO772886	AO772886
266	17	6.0	485	10	AV637943	AV637943	AV637943	339	17	6.0	542	10	AV388053	AV388053
267	17	6.0	487	10	AV644626	AV644626	AV644626	340	17	6.0	542	10	BE022502	BE022502
268	17	6.0	488	10	AV634179	AV634179	AV634179	341	17	6.0	543	10	AV386908	AV386908
269	17	6.0	488	10	AV635942	AV635942	AV635942	342	17	6.0	544	10	AV620198	AV620198
270	17	6.0	489	10	AV633200	AV633200	AV633200	343	17	6.0	544	10	AV619621	AV619621
271	17	6.0	491	10	AV621196	AV621196	AV621196	344	17	6.0	544	11	BF076541	BF076541
272	17	6.0	492	10	AV634408	AV634408	AV634408	345	17	6.0	547	10	AV622992	AV622992
273	17	6.0	492	10	AV643335	AV643335	AV643335	346	17	6.0	547	10	AV627803	AV627803
274	17	6.0	493	10	AI224387	AI224387	AI224387	347	17	6.0	550	13	AQ540692	AQ540692
275	17	6.0	493	10	AV634386	AV634386	AV634386	348	17	6.0	551	10	AV392141	AV392141
276	17	6.0	493	10	AV642988	AV642988	AV642988	349	17	6.0	552	10	AA485746	AA485746
277	17	6.0	495	10	AV623067	AV623067	AV623067	350	17	6.0	554	11	BG362713	BG362713
278	17	6.0	495	10	AV638799	AV638799	AV638799	351	17	6.0	555	10	AV387877	AV387877
279	17	6.0	495	10	AW707651	AW707651	AW707651	352	17	6.0	556	10	AV390143	AV390143
280	17	6.0	496	10	AW757875	AW757875	AW757875	353	17	6.0	557	10	AI404323	AI404323
281	17	6.0	497	10	AV636793	AV636793	AV636793	354	17	6.0	558	10	AW291702	AW291702
282	17	6.0	498	10	AW773002	AW773002	AW773002	355	17	6.0	561	10	AV640773	AV640773
283	17	6.0	499	10	AV623365	AV623365	AV623365	356	17	6.0	562	10	AI822125	AI822125
C 284	17	6.0	500	13	AQ434023	AQ434023	AQ434023	C 357	17	6.0	564	13	AQ497594	AQ497594
285	17	6.0	501	10	AV640577	AV640577	AV640577	358	17	6.0	565	10	AV388971	AV388971
286	17	6.0	502	10	AV622508	AV622508	AV622508	C 359	17	6.0	565	11	BG376012	BG376012
287	17	6.0	502	10	AV635995	AV635995	AV635995	360	17	6.0	566	10	BE227856	BE227856
288	17	6.0	502	11	BC789703	BC789703	BC789703	361	17	6.0	570	10	AM459511	AM459511
C 289	17	6.0	502	13	AZ178275	AZ178275	AZ178275	362	17	6.0	571	11	BF567564	BF567564
290	17	6.0	504	10	AV627813	AV627813	AV627813	363	17	6.0	573	10	BE724406	BE724406
291	17	6.0	505	10	AV623066	AV623066	AV623066	364	17	6.0	575	13	AQ267314	AQ267314
C 292	17	6.0	506	10	AJ393779	AJ393779	AJ393779	C 365	17	6.0	578	13	AZ368919	AZ368919
293	17	6.0	507	10	AV633785	AV633785	AV633785	366	17	6.0	579	11	BG709633	BG709633
294	17	6.0	507	10	AV637571	AV637571	AV637571	367	17	6.0	582	10	AW721469	AW721469
295	17	6.0	507	10	AV642248	AV642248	AV642248	368	17	6.0	583	10	AW391351	AW391351
296	17	6.0	508	10	AV624230	AV624230	AV624230	C 369	17	6.0	589	11	BG237974	BG237974
297	17	6.0	508	10	AV639841	AV639841	AV639841	370	17	6.0	590	10	AI883110	AI883110
298	17	6.0	510	10	AV644281	AV644281	AV644281	371	17	6.0	590	10	AW707430	AW707430
C 299	17	6.0	510	11	BI291978	BI291978	BI291978	372	17	6.0	593	10	BE452983	BE452983
300	17	6.0	510	13	AZ639069	AZ639069	AZ639069	373	17	6.0	598	10	AI062405	AI062405
301	17	6.0	511	10	AV633717	AV633717	AV633717	374	17	6.0	601	10	AI485410	AI485410
302	17	6.0	512	10	AV623649	AV623649	AV623649	375	17	6.0	606	11	BI066581	BI066581
303	17	6.0	513	10	BE129170	BE129170	BE129170	376	17	6.0	608	10	AW757590	AW757590
304	17	6.0	514	10	AV641994	AV641994	AV641994	377	17	6.0	609	10	BE726134	BE726134

C 524	16	5.7	311	10	AI033408	cx03f02.s	597	16	5.7	388	10	AI084601	AI084601 ox63ff10.s
525	16	5.7	311	10	BB461452	BB461452	598	16	5.7	388	11	H00239	H00239 yj22h06.r1
526	16	5.7	311	11	BG410655	B47051E11	599	16	5.7	389	10	AA690154	AA690154 vt82b05.s
C 527	16	5.7	312	10	AM377679	PMO-CT023	600	16	5.7	389	11	W70361	W70361 me19g04.r1
528	16	5.7	314	10	AI202731	q145602.x	601	16	5.7	389	13	AO305084	AO305084 HS.2017.A
C 529	16	5.7	314	10	AM135912	UI-H-B11-	602	16	5.7	390	10	AA034007	AA034007 z105b06.s
C 530	16	5.7	317	10	AM880387	QV0-CT003	603	16	5.7	390	13	AA283493	AA283493 2M0097D1.s
531	16	5.7	319	10	AI978559	wq72a05.x	604	16	5.7	392	10	AM260776	AM260776 um87c08.y
C 532	16	5.7	319	11	BF016480	uy37c03.y	605	16	5.7	392	10	AM647355	AM647355 EST325898
C 533	16	5.7	322	10	AA328950	EST32776	606	16	5.7	393	11	BF991306	BF991306 CM0-GN016
534	16	5.7	322	11	BF713329	MI-P-C02-a	607	16	5.7	395	10	AI286530	AI286530 ui85h10.y
C 535	16	5.7	323	13	A2357629	IM0099M14	608	16	5.7	397	10	AM571206	AM571206 ra20b09.y
536	16	5.7	328	11	R10661	yf13f10.r1	609	16	5.7	400	10	AM469747	AM469747 hd34b06.x
C 537	16	5.7	329	10	AA149598	AA149598 zc003h06.s	610	16	5.7	400	11	BI019460	BI019460 CM3-MT019
C 538	16	5.7	330	10	AAV657124	AV657124	611	16	5.7	401	10	AB029125	AB029125 AB029125
C 539	16	5.7	330	11	BG982509	IL5-CN006	612	16	5.7	403	10	AA912996	AA912996 o128c10.x
C 540	16	5.7	330	11	BF448007	BF448007 nae3b01.	613	16	5.7	403	10	AI216276	AI216276 q976a07.s
541	16	5.7	337	11	R50154	R50154 yj61d03.s1	614	16	5.7	406	10	AA890713	AA890713 ak13c02.s
542	16	5.7	337	11	BF918599	CM0-MT013	615	16	5.7	406	11	RA8766	RA8766 yj69a04.s1
543	16	5.7	338	11	W84590	W84590 zd91c06.r1	616	16	5.7	406	11	BG232096	BG232096 na132f09.
C 544	16	5.7	338	13	A2805655	A2805655 2M0067111	617	16	5.7	407	10	AM264576	AM264576 xr05b04.x
C 545	16	5.7	338	13	A2892075	RPC1-24-1	618	16	5.7	410	10	AI657034	AI657034 tt49c10.x
546	16	5.7	339	13	AQ007511	CIT-HSP-2	619	16	5.7	410	10	AA435083	AA435083 vf41g08.r
C 547	16	5.7	341	10	AA777540	A777540 z194b09.s	620	16	5.7	410	10	AA455861	AA455861 aa01a12.s
C 548	16	5.7	341	13	AZ035619	RPC1-23-3	621	16	5.7	411	10	AI571198	AI571198 tm43c10.x
C 549	16	5.7	344	10	AA036443	AA036443 m156b07.r	622	16	5.7	411	11	BF714591	BF714591 mas01g06.
C 550	16	5.7	347	10	AW934982	CM2-DT000	623	16	5.7	411	13	AO148771	AO148771 HS.3136.A
551	16	5.7	347	11	W87810	W87810 zh66g07.s1	624	16	5.7	412	10	AA620909	AA620909 a196a11.s
552	16	5.7	348	10	AA467715	AA467715 ve02g09.r	625	16	5.7	412	13	AA055928	AA055928 HS.5151.B
553	16	5.7	348	11	R57510	R57510 F3312 fetal	626	16	5.7	413	11	BI023428	BI023428 CM4-MT024
C 554	16	5.7	349	11	F10673	F10673 HSC31D032.n	627	16	5.7	413	11	BF410972	BF410972 UI-R-CN0-
555	16	5.7	350	10	AM226737	AM226737 um60c05.y	628	16	5.7	414	10	AA494430	AA494430 ne28c09.s
556	16	5.7	350	11	BI096719	BI096719 949013A06	629	16	5.7	416	10	AI206971	AI206971 q130h02.x
C 557	16	5.7	350	13	AO165665	AO165665 HS.3065.B	630	16	5.7	416	11	BF812734	BF812734 RC3-CI019
C 558	16	5.7	352	10	AI165367	AI165367 A082p32u	631	16	5.7	417	11	BF060732	BF060732 7j60a03.x
559	16	5.7	352	10	AI521288	AI521288 L104d10.x	632	16	5.7	417	13	BH077807	BH077807 RPC1-24-2
C 560	16	5.7	353	11	H45525	H45525 y112g12.r1	633	16	5.7	418	10	AI590135	AI590135 fo12e03.x
C 561	16	5.7	355	10	AV656653	AV656653	634	16	5.7	419	10	AV664657	AV664657 AV664657
C 562	16	5.7	356	10	AM663370	h167b03.x	635	16	5.7	419	10	AA207589	AA207589 mv81c11.r
563	16	5.7	357	11	BF904728	BF904728 IL5-MT025	636	16	5.7	420	10	AU014637	AU014637 A014637
C 564	16	5.7	357	10	AV658648	AV658648	637	16	5.7	420	10	AI283188	AI283188 qk49b01.x
565	16	5.7	359	10	AM141319	AM141319 EST291360	638	16	5.7	420	10	AM024579	AM024579 wu77f11.x
C 566	16	5.7	360	11	BI045797	MR3-FN020	639	16	5.7	421	11	BF721024	BF721024 mab61b03.
C 567	16	5.7	362	10	AA298434	AA298434 EST114198	640	16	5.7	422	10	AA813723	AA813723 a171e07.s
C 568	16	5.7	363	10	AA722476	AA722476 z991d10.s	641	16	5.7	422	10	BE044126	BE044126 ho37g11.x
C 569	16	5.7	363	13	A2744400	RPC1-24-7	642	16	5.7	422	10	BE652934	BE652934 UI-M-AM1-
C 570	16	5.7	364	10	AA812705	AA812705 aj31f03.s	643	16	5.7	423	10	AI671694	AI671694 ty26h01.x
571	16	5.7	365	10	BE327217	BE327217 hw08g08.x	644	16	5.7	423	10	AI880832	AI880832 at11c06.x
572	16	5.7	365	11	W76709	W76709 me73c09.r1	645	16	5.7	425	10	AI121590	AI121590 uc43f06.r
573	16	5.7	366	11	R72122	R72122 yj87c03.r1	646	16	5.7	425	13	AQ459073	AQ459073 HS.5068.B
574	16	5.7	368	13	CNS00017	AL085225 Arabidops	647	16	5.7	426	11	BI010176	BI010176 CMI-LT004
575	16	5.7	369	11	N67484	N67484 za07c12.s1	648	16	5.7	427	11	R23862	R23862 y448e08.r1
C 576	16	5.7	370	10	AM301298	AM301298 xs70h11.x	649	16	5.7	427	11	BF108490	BF108490 7n58f02.x
C 577	16	5.7	371	10	AA967777	uh04f10.r	650	16	5.7	428	13	BH068135	BH068135 RPC1-24-3
578	16	5.7	371	13	AQ003197	RPC111-27	651	16	5.7	429	10	AA450743	AA450743 vf78c11.r
C 579	16	5.7	372	11	BI179282	EST520227	652	16	5.7	429	11	TS5431	TS5431 ya69a06.s1
C 580	16	5.7	374	10	AI285095	AI285095 qk70b06.x	653	16	5.7	430	10	AA636285	AA636285 vg76b12.s
581	16	5.7	374	10	AU015660	AU015660	654	16	5.7	432	10	BE632677	BE632677 uv64a01.y
C 582	16	5.7	374	10	AM772476	AM772476 hm75a03.x	655	16	5.7	432	11	N25863	N25863 yx22h08.r1
C 583	16	5.7	374	10	BE047999	BE047999 tz45c09.y	656	16	5.7	432	13	BA5321	BA5321 HS-1061-A1-
584	16	5.7	376	10	AA560665	AA560665 v171a09.r	657	16	5.7	433	11	BF852364	BF852364 MR3-EN009
585	16	5.7	376	11	M78469	M78469 EST00617.Fe	658	16	5.7	434	10	AI116907	AI116907 ue21h07.y
C 586	16	5.7	378	10	AA189182	AA189182 mu46a09.r	659	16	5.7	435	10	AM988284	AM988284 ug05d03.y
C 587	16	5.7	378	11	BI046562	BI046562 MX3-FN020	660	16	5.7	435	10	AM988316	AM988316 ug06d03.y
588	16	5.7	379	10	AM089827	AM089827 xdl1g10.x	661	16	5.7	435	13	AO508819	AO508819 RPC1-11*2
589	16	5.7	380	11	W36906	W36906 mb77h09.r1	662	16	5.7	436	10	AA197398	AA197398 mu17h06.r
590	16	5.7	381	10	AM320813	u023h05.y	663	16	5.7	437	11	BF011586	BF011586 us35a07.y
C 591	16	5.7	381	11	H52938	H52938 yq76f11.r1	664	16	5.7	438	11	AA188311	AA188311 zp71c09.r
C 592	16	5.7	382	10	AM984323	PM3-HN001	665	16	5.7	438	11	T32956	T32956 EST56060.Hu
C 593	16	5.7	383	10	AJ285112	AJ285112 AA3B-AAF-	666	16	5.7	439	11	BF519624	BF519624 EST457088
594	16	5.7	385	10	AL134336	AL134336 DKF2p5470	667	16	5.7	440	11	BE558231	BE558231 UI-R-A1-e
595	16	5.7	385	11	H73470	H73470 ys12c08.r1	668	16	5.7	440	13	AZ110482	AZ110482 RPC1-23-8
596	16	5.7	387	10	AA823591	AA823591 vp41h09.r	669	16	5.7	441	10	AA423608	AA423608 ve77b04.r

C 816	16	5.7	523	13	AO507201	RPCI -11-3	AO507201	RPCI -11-3	889	16	5.7	572	13	AO802915	HS_3160..A	AO802915	HS_3160..A
C 817	16	5.7	524	10	BB283667	BB283667	BB283667	BB283667	C 890	16	5.7	573	11	BS396263	602459049	BS396263	602459049
C 818	16	5.7	525	10	AA606788	vm90e10..r	AA606788	vm90e10..r	C 891	16	5.7	573	13	Az217106	Shneared D	Az217106	Shneared D
C 819	16	5.7	527	11	B1319967	B1319967	B1319967	B1319967	C 892	16	5.7	574	10	AW963277	ESTS73530	AW963277	ESTS73530
C 820	16	5.7	528	10	A1459235	tk11c04..x	A1459235	tk11c04..x	C 893	16	5.7	574	10	AA619825	AA619825	AA619825	v161d05..s
C 821	16	5.7	528	10	AW249439	AW249439	AW249439	2819339..s	C 894	16	5.7	575	10	AA473817	vq92e06..r	AA473817	vq92e06..r
C 822	16	5.7	529	10	AW026837	AW026837	AW026837	w864f11..x	C 895	16	5.7	575	11	TF15739	TF15739	TF15739	TF18185..Inta
C 823	16	5.7	529	10	BE039727	BE039727	BE039727	OC07A08 O	C 896	16	5.7	575	11	BF199133	BF199133	BF199133	249212 MA
C 824	16	5.7	529	10	BE398108	BE398108	BE398108	601290328	C 897	16	5.7	576	10	AA530706	AA530706	AA530706	v150d05..r
C 825	16	5.7	529	11	B1046568	B1046568	B1046568	MR3-FN020	C 898	16	5.7	577	11	BC064985	BC064985	BC064985	H3026B01-
C 826	16	5.7	529	13	Az317863	Az317863	IM0036H05	Az317863	C 899	16	5.7	578	10	A1990526	A1990526	A1990526	w840908.x
C 827	16	5.7	531	13	Az155081	SP_0013-A	Az155081	SP_0013-A	C 900	16	5.7	579	10	AA839596	AA839596	AA839596	vw97b11..r
C 828	16	5.7	531	13	BH095792	RPCI -24-3	BH095792	RPCI -24-3	C 901	16	5.7	579	10	BE739531	BE739531	BE739531	601556467
C 829	16	5.7	531	13	AO436133	HS_5049..B	AO436133	HS_5049..B	C 902	16	5.7	579	13	AO442673	AO442673	AO442673	HS_5116..B
C 830	16	5.7	533	10	A1519608	LD39253..5	A1519608	LD39253..5	C 903	16	5.7	581	11	B1341563	B1341563	B1341563	369031 MA
C 831	16	5.7	534	11	BF590338	BF590338	nab23b11	BF590338	C 904	16	5.7	581	13	AO600338	AO600338	AO600338	HS_5360..B
C 832	16	5.7	534	13	B56718	CTT-HSP-200	B56718	CTT-HSP-200	C 905	16	5.7	585	11	BF703480	BF703480	BF703480	MI-P-H1..g
C 833	16	5.7	535	10	AV391170	AV391170	AV391170	AV391170	C 906	16	5.7	585	13	Az736910	Az736910	Az736910	RPCI -24-7
C 834	16	5.7	535	13	Az323184	IM0044A22	Az323184	IM0044A22	C 907	16	5.7	586	10	AM386287	AM386287	AM386287	CM1-PT001
C 835	16	5.7	536	10	A1880840	at11d06..x	A1880840	at11d06..x	C 908	16	5.7	586	11	BG150559	BG150559	BG150559	7X02f07.x
C 836	16	5.7	536	11	H16417	ym23c01..r1	H16417	ym23c01..r1	C 909	16	5.7	587	10	AL588527	AL588527	AL588527	AL588527
C 837	16	5.7	536	13	AO808457	HS_5298..B	AO808457	HS_5298..B	C 910	16	5.7	587	11	BC088219	BC088219	BC088219	H3150C06
C 838	16	5.7	537	10	BE102320	UI-R-BT1-	BE102320	UI-R-BT1-	C 911	16	5.7	587	13	Az304566	Az304566	Az304566	

C 962	16	5.7	628	13	A0534300	A0534300	RPCI-11-3	A0534300
C 963	16	5.7	629	11	BF347251	BF347251	6020211963	BF347251
C 964	16	5.7	630	13	A2874257	A2874257	2M0168M1	A2874257
C 965	16	5.7	631	13	A2574141	A2574141	325PVC09	A2574141
C 966	16	5.7	632	10	AW536533	AW536533	G0105F04	AW536533
C 967	16	5.7	633	10	A1999998	A1999998	614008B05	A1999998
C 968	16	5.7	634	10	BF268416	BF268416	GA_EB000	BF268416
C 969	16	5.7	637	10	A1486785	A1486785	ESM2451075	A1486785
C 970	16	5.7	637	13	A2109873	A2109873	RPCI-23-8	A2109873
C 971	16	5.7	638	13	A2983763	A2983763	2M0265102	A2983763
C 972	16	5.7	640	10	A1982153	A1982153	pat_PK007	A1982153
C 973	16	5.7	640	11	BG848217	BG848217	1024021B0	BG848217
C 974	16	5.7	641	11	BG818401	BG818401	602279945	BG818401
C 975	16	5.7	645	11	BE377494	BE377494	601228972	BE377494
C 976	16	5.7	645	11	BP530828	BP530828	602072191	BP530828
C 977	16	5.7	648	11	B117064	B117064	602867855	B117064
C 978	16	5.7	649	13	A2791487	A2791487	2M0041K24	A2791487
C 979	16	5.7	650	11	BG921098	BG921098	602825366	BG921098
C 980	16	5.7	650	11	BG964298	BG964298	602831906	BG964298
C 981	16	5.7	653	11	B1150688	B1150688	602915066	B1150688
C 982	16	5.7	654	11	B1231369	B1231369	RE21427.5	B1231369
C 983	16	5.7	658	11	BG668029	BG668029	DKABTFF03	BG668029
C 984	16	5.7	659	11	B1249398	B1249398	602995718	B1249398
C 985	16	5.7	659	13	A0157326	A0157326	npxb00099	A0157326
C 986	16	5.7	662	11	BF141827	BF141827	601791301	BF141827
C 987	16	5.7	663	10	AV703910	AV703910	AV703910	AV703910
C 988	16	5.7	663	10	BE447894	BE447894	uT41502_Y	BE447894
C 989	16	5.7	664	10	BE258365	BE258365	6011148645	BE258365
C 990	16	5.7	665	13	A2580686	A2580686	1M0369N14	A2580686
C 991	16	5.7	666	11	BG873570	BG873570	602792178	BG873570
C 992	16	5.7	666	11	BG914818	BG914818	602813626	BG914818
C 993	16	5.7	666	11	B1148357	B1148357	602914378	B1148357
C 994	16	5.7	671	11	B1275681	B1275681	UI-R-CW0	B1275681
C 995	16	5.7	671	13	A2870574	A2870574	2M0183B11	A2870574
C 996	16	5.7	672	10	AL562198	AL562198	AL562198	AL562198
C 997	16	5.7	674	10	AW744485	AW744485	ur33304_Y	AW744485
C 998	16	5.7	675	11	BG859606	BG859606	1024056D0	BG859606
C 999	16	5.7	675	11	BE898180	BE898180	601435528	BE898180
C1000	16	5.7	676	10	AU170159	AU170159	AU170159	AU170159

ALIGNMENTS

RESULT	1
LOCUS	H34451
DEFINITION	H34451 319 bp mRNA EST 13-MAR-1998
ACCESSION	EST11406 Rat PC-12 cells, NGF-treated (9 days) Rattus sp. cDNA 5'
VERSION	H34451
KEYWORDS	end similar to Glucose transporter, mRNA sequence.
SOURCE	H34451.1 GI:979868
ORGANISM	EST.
	Rattus sp.
	Rattus sp.
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
	Rattus.
	1 (bases 1 to 319)
REFERENCE	Lee N.H., Weinstein K.G., Kirkness E.F., Earle-Hughes J.A., Fulne
AUTHORS	, R.A., Mammas S., Glodok A., Gocayne J.D., Adams M.D., Kerlavage
	, A.R., Fraser C.M. and Venter J.C.
	Comparative expressed-sequence-tag analysis of differential gene
	expression profiles in PC-12 cells before and after nerve growth
	factor treatment
TITLE	Proc. Natl. Acad. Sci. U.S.A. 92, 8303-8307 (1995)
JOURNAL	95396786
MEDLINE	Contact: Lee, NH
COMMENT	The Institute for Genomic Research
	9712, Medical Center Drive, Rockville, MD 20850, USA
	Tel: (301)-838-3529
	Fax: (301)-838-0208
	Email: nhlee@igf.org
	For clone availability please contact the TIGR Database

FEATURES		Seq primer: M13 Reverse.	
source	Location/Qualifiers		
	1. 319		
	/organism="Rattus sp.."		
	/db_xref="ATCC (Inhost):2005510"		
	/db_xref="taxon:10118"		
	/clone_lib="Rat PC-12 cells, NGF-treated (9 days)"		
	/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; poly(A)+ RNA was purified from 9-day NGF treated PC12 cells. cDNA was constructed using an oligo-dT primer and directionally cloned using the Lambda Zap II Vector kit by Stratagene"		
BASE COUNT	48 a 102 c 80 g 88 t 1 others		
ORIGIN			
Query Match	63.8%; Score 180; DB 11; Length 319;		
Best Local Similarity	99.3%; Pred. No. 2, 2e-79;		
Matches 280; Conservative	0; Mismatches 2; Indels 0; Gaps 0		
0Y	1 gaaacctgcatggttcaacctggagggctgagcctgctgctgctgctgataagcagatgctccttc 60		
Db	19 GAGCCTGCTGAGTGTTCACCTGGGGCTGGCTGGCTGGCTGTAAGCGCATGTGCCCTCTTC 78		
0Y	61 atcgcgtgttttgcagtaagctgctggaggaaccatccctctgctcctcatatgcaagatcttc 120		
Db	79 ATCGGCTGGTTTGGCAGTAGGCTGGGGAGCCCATCCCTGGCTCTCATGTACAGAGATCTTC 138		
0Y	121 cctctgcacatcaaggtgtgtgctaccggcgctctgtgtctctcaaccaactggttcatgtgcc 180		
Db	139 CCTGTGCACATCAAGAGGTGTGGCTACCGGCGTCTGTTCCTCACCAACTGTTCATGAGCC 198		
0Y	181 ttctctgtgtgaccaaagagtttaacagcatcatgtgagatctctcaagacctagggcgcttc 240		
Db	199 TTTTGTGGTGACCAAGAGATTGTAAACGATCATGAGAGATCTTCAGACCTCAGCGCGCTTC 258		
0Y	241 tggctcacacgctgctctgtatctctcaagcgctcttttaag 282		
Db	259 TGGCTACCGGCTGCTGTGTATCTTCACAGCTCTTTTACG 300		

RESULT	2
BG079217	
LOCUS	
DEFINITION	BG079217 595 bp mRNA EST 26-JAN-2001
ACCESSION	H3037C03-5 NIA Mouse 15K cDNA Clone Set Mus musculus CDNA clone
VERSION	H3037C03 5', mRNA sequence.
KEYWORDS	BG079217
SOURCE	BG079217.1 GI:12561785
ORGANISM	EST.
TITLE	house mouse.
JOURNAL	Mus musculus
COMMENT	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 595) Karagi,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka, T.S., Carter,M.G. and Ko,M.S.H. Verification and initial annotation of NIA mouse 15K cDNA clone set Unpublished (2001)
REFERENCE	Other_ESTs: H3037C03-3
AUTHORS	Contact: George J. Karagi
	Laboratory of Genetics
	National Institute on Aging/National Institutes of Health
	333 Cassell Drive, Suite 400, Baltimore, MD 21224-6820, USA
	Email: cdnaelgsun.grc.nia.nih.gov
	This clone set has been freely distributed to the community. Please
	visit http://lgsun.grc.nia.nih.gov/CDNA/15k.html for details.
	Plate: H3037 row: C column: 03
	Seq primer: -21M13 Reverse
	High quality sequence stop: 595
	POLYA-No.
FEATURES	
source	Location/Qualifiers
	1..595

```

/organism="Mus musculus"
/strain="C57Bl/6J"
/db_xref="taxon:10090"
/clone_1lb="H3037C03"
/clone_1lb="NIA Mouse 15K cDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/notes="Vector: pSPORT1. Site 1: SalI; Site 2: NotI. This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT) Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."
Mol Genet 7: 1967-1978."
BASE COUNT      95 a      177 c      182 g      141 t
ORIGIN

```

```

Query Match      17.7%: Score 50; DB 11; Length 595;
Best Local Similarity 100.0%: Pred. No. 1.8e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 79 ggcctggagaccatccctgcctccatgacagatctccctcga 128
|||||
Db 516 GCGTGGGAGCCATCCCTGCTCATGTGCAGATCTTCCCTGTGCA 565

```

```

RESULT 3
LOCUS      BF120840      823 bp      mRNA      EST      24-OCT-2000
DEFINITION 601757442F1 NCL_CGAP_Mam5 Mus musculus cDNA clone IMAGE:3986651 5',
VERSION    BF120840
KEYWORDS   BF120840.1 GI:10959880
SOURCE     EST.
ORGANISM   house mouse.
            Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 823)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabs-remail.nih.gov
            Tissue Procurement: Lotmar Hennighausen Ph.D., Robin Humphreys
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: LHAM9192 row: d column: 12
            High quality sequence start: 4
            High quality sequence stop: 644.
            Location/Qualifiers
                1..823
                /organism="Mus musculus"
                /strain="C57/B6"

```

```

FEATURES
source

```

```

/db_xref="taxon:10090"
/clone="IMAGE:3986651"
/clone_1lb="NCL CGAP Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; site_1: SalI;
site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lotmar Hennighausen/Robin Humphreys,
NIH"
BASE COUNT      150 a      226 c      257 g      190 t
ORIGIN

```

```

Query Match      17.7%: Score 50; DB 11; Length 823;
Best Local Similarity 100.0%: Pred. No. 1.8e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 79 ggcctggagaccatccctgcctccatgacagatctccctcga 128
|||||
Db 282 GCGTGGGAGCCATCCCTGCTCATGTGCAGATCTTCCCTGTGCA 331

```

```

RESULT 4
LOCUS      BF140667      917 bp      mRNA      EST      24-OCT-2000
DEFINITION 601786917F1 NCL_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4014605 5',
VERSION    BF140667
KEYWORDS   BF140667.1 GI:10979707
SOURCE     EST.
ORGANISM   house mouse.
            Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 917)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabs-remail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: LHAM9259 row: a column: 06
            High quality sequence stop: 605.
            Location/Qualifiers
                1..917
                /organism="Mus musculus"
                /strain="C57/B6"
                /db_xref="taxon:10090"
                /clone="IMAGE:4014605"
                /clone_1lb="NCL CGAP Lu30"
                /tissue_type="tumor, metastatic to mammary"
                /lab_host="DH10B"
                /notes="Organ: lung; Vector: pCMV-SPORT6; site_1: NotI;
                site_2: SalI; transgenic model WWT-1, expression driven by
                MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dT.
                Library constructed by Life Technologies.
                Investigator providing samples: Gilbert Smith, NIH"

```

```

BASE COUNT      163 a      263 c      277 g      214 t
ORIGIN

```

```

Query Match      17.0%: Score 48; DB 11; Length 917;
Best Local Similarity 100.0%: Pred. No. 1.8e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 131 tcaagggtggtgctaccgagcgtgctgctccccaactggtcatg 178
 |||
 Db 559 TCAAGGCTGTGCTACCGCGCTGTCTGTCTCACCACACTGGTTCATGC 606

RESULT 5
 LOCUS BF774206 440 bp mRNA EST 25-APR-2001
 DEFINITION B283798 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION BF774206
 VERSION BF774206.1 GI:12122106
 KEYWORDS EST.
 SOURCE
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 440)
 AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,
 G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-Mokom,C.G.,
 Petrea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
 Keefe,J.W.
 TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
 JOURNAL Genome Res. 11 (4), 626-630 (2001)
 MEDLINE 21180013
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smithemail.marc.usda.gov
 Single pass sequencing. Bases called and alt-trimmed with phred
 v0.980904.e. Vector identified by cross-match with the -minscore 18
 and -minmatch 12 options.
 PCR primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGTCACGACG
 Plate: 83 row: H column: 1
 Seq primer: ATTAGCTGACACTATAG.
 FEATURES
 Location/Qualifiers
 Source 1..440
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 3BOV"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 library made from pooled tissue from marrow, alveolar
 macrophage, ovary, fetal semitendinosus muscle, and fetal
 longissimus muscle."
 BASE COUNT 65 a 141 c 135 g 99 t
 ORIGIN

Query Match 15.6%; Score 44; DB 11; Length 440;
 Best Local Similarity 100.0%; Pred. No. 1.7e-11;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 137 gtcctcaccactggttcacgtccttctggtgaccaagaagtc 200
 |||
 Db 385 GTCTTCACCACTGCTTCATGCGCTTCTGCTGACCAAGAGTT 428

RESULT 6
 LOCUS BF742266 327 bp mRNA EST 10-JAN-2001
 DEFINITION RCI-BTN0409-021000-012-e12 BTN0409 Homo sapiens cDNA, mRNA
 ACCESSION BF742266
 VERSION BF742266.1 GI:12069046
 KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 327)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W.Jr., Zaigo,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC1&t2=RC1-BTN0409>
 -021000-012-e12&t3=2000-10-02&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 11
 High quality sequence stop: 327.
 FEATURES
 Location/Qualifiers
 Source 1..327
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="BTN0409"
 /dev_stage="Adult"
 /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
 SmaI; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No. 196
 716 - Ludwig Institute for Cancer Research) profiles
 into the puc 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
 BASE COUNT 82 a 91 c 108 g 46 t
 ORIGIN

Query Match 14.5%; Score 41; DB 11; Length 327;
 Best Local Similarity 100.0%; Pred. No. 5.4e-10;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 88 cccatccctgctcctcatgtcagaagatcttcctcgca 128
 |||
 Db 195 CCCATCCCTGCTCTCATGTGACAGATCTTCCTCTGCA 155

RESULT 7
 LOCUS AA627408 358 bp mRNA EST 31-OCT-1997
 DEFINITION nq49d04.S1 NCI CGAP Colo Homo sapiens cDNA clone IMAGE:1147207 3'
 similar to TR:G1209756 G1209756 INTEGRAL MEMBRANE PROTEIN.;; mRNA
 ACCESSION AA627408
 VERSION AA627408
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 358)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Ian Kirsch, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bdrrp/image/image.html
 Insert length: 507 Std Error: 0.00
 Seq primer: -40ml3 fwd. RT from Amersham
 High quality sequence stop: 349.
FEATURES
 source
 1. 358
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:1147207"
 /clone_lib="NCI-CGAP-Co10"
 /tissue_type="colon tumor RER+."
 /lab_host="DH10B"
 /note="Organ: colon; Vector: pFT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from RER+ colon tumor, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pFT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonalido (N-Soares4)."
BASE COUNT
 ORIGIN
 87 a 104 c 101 g 66 t
 Query Match 14.5%: Score 41; DB 10; Length 358;
 Best Local Similarity 100.0%: Pred. No. 5.4e-10;
 Matches 41: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 351 CCCATCCCTGCTCTCATGTCTCAGACATCTCCCTCTGCA 311
 ||||||||||||||||||||||||||||||||||||||||
 88 cccatccctgctctcatgtcagagatctccctctgca 128
 ||||||||||||||||||||||||||||||||||||||||
RESULT 8
LOCUS AM248655 496 bp mRNA EST 07-JAN-2000
DEFINITION 2820759.3prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820759 3', mRNA sequence.
ACCESSION AM248655
VERSION AM248655.1 GI:6591648
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
REFERENCE 1 (bases 1 to 496)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Other_ESTs: 2820759.5prime
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: DCTD/DRP cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Berkeley MGC Sequencing Project Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bdrrp/image/image.html Base Calling / Quality Scores: PHRED from University of Washington Genome Center
 Trimming: cross-match from University of Washington Genome Center
 PHRAP suite. Poly-T identification: palMatch.pl from Berkeley
 Drosophila genome project. University of Washington Genome Center:
 http://www.genome.washington.edu/polyadenylation: Based upon the

presence of a XhoI site followed by a run of 14 or more T residues at the beginning of the sequence, this cDNA insert was polyadenylated.
 Plate: LLCM5 row: A column: 16
 High quality sequence stop: 415.
FEATURES
 source
 1. 496
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:2820759"
 /clone_lib="NIH_MGC_7"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pORT67; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(6). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT
 ORIGIN
 116 a 145 c 142 g 92 t 1 others
 Query Match 14.5%: Score 41; DB 10; Length 496;
 Best Local Similarity 100.0%: Pred. No. 5.5e-10;
 Matches 41: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 368 CCCATCCCTGCTCTCATGTCTCAGACATCTCCCTCTGCA 328
 ||||||||||||||||||||||||||||||||||||||||
 88 cccatccctgctctcatgtcagagatctccctctgca 128
 ||||||||||||||||||||||||||||||||||||||||
RESULT 9
LOCUS BG290178 712 bp mRNA EST 21-FEB-2001
DEFINITION 602385265F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4514200 5', mRNA sequence.
ACCESSION BG290178
VERSION BG290178.1 GI:13046711
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
REFERENCE 1 (bases 1 to 712)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: LLAM10402 row: a column: 17
 High quality sequence stop: 707.
FEATURES
 source
 1. 712
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:4514200"
 /clone_lib="NIH_MGC_93"
 /tissue_type="transitional cell papilloma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.7 kb. Library enriched for

full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 117 a 221 c 211 g 163 t

Query Match 14.5%; Score 41; DB 11; Length 712;
 Best Local Similarity 100.0%; Pred. No. 5,6e-10;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 88 cccatccctgctctcatgctcagagatctccctctgca 128
 ||||||||||||||||||||||||||||||||||||||||
 Db 74 cccatccctgctctcatgctcagagatctccctctgca 114

RESULT 10

BE910478 934 bp mRNA EST

LOCUS 601501005F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3903048 5',

DEFINITION mRNA sequence.

ACCESSION BE910478

VERSION BE910478.1

KEYWORDS GI:10407108

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: LCM1681 row: P column: 12
 High quality sequence stop: 745.

FEATURES

source 1..949
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="4844411"
 /clone_lib="NIH_MGC_43"
 /tissue_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pORF7; Site: 1; XhoI; Site: 2;
 EcoRI; CDNA made by oligo-dt priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library. |"

REFERENCE 1 (bases 1 to 934)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: LCM9706 row: 1 column: 01
 High quality sequence stop: 710.

FEATURES

source 1..934
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="3903048"
 /clone_lib="NIH_MGC_70"
 /tissue_type="epitheloid carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: pancreas; Vector: pCMV-SPORT6; Site: 1; NotI;
 Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.1 kb. Library constructed by Life
 Technologies."

BASE COUNT 143 a 291 c 287 g 213 t

ORIGIN

Query Match 14.5%; Score 41; DB 11; Length 934;
 Best Local Similarity 100.0%; Pred. No. 5,7e-10;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 88 cccatccctgctctcatgctcagagatctccctctgca 128
 ||||||||||||||||||||||||||||||||||||||||
 Db 462 cccatccctgctctcatgctcagagatctccctctgca 502

RESULT 11

BE749509 949 bp mRNA EST

LOCUS 602707648F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4844411 5',

DEFINITION mRNA sequence.

ACCESSION BE749509

VERSION BE749509

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Other_ESTs: 2820759, 3prime
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DPF CDNA Library Preparation: Ling
 Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E.
 Consortium (LNLN) DNA Sequencing by: Berkeley MGC sequencing
 project Clone distribution: MGC clone distribution information can

VERSION BE749509.1 GI:14060162
 EST.
 KEYWORDS human.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 949)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: LCM1681 row: P column: 12
 High quality sequence stop: 745.

FEATURES

source 1..949
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="4844411"
 /clone_lib="NIH_MGC_43"
 /tissue_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pORF7; Site: 1; XhoI; Site: 2;
 EcoRI; CDNA made by oligo-dt priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library. |"

BASE COUNT 160 a 302 c 264 g 223 t

ORIGIN

Query Match 14.5%; Score 41; DB 11; Length 949;
 Best Local Similarity 100.0%; Pred. No. 5,7e-10;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 88 cccatccctgctctcatgctcagagatctccctctgca 128
 ||||||||||||||||||||||||||||||||||||||||
 Db 167 cccatccctgctctcatgctcagagatctccctctgca 207

RESULT 12

AM249090 480 bp mRNA EST

LOCUS 2820759.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820759 5',

DEFINITION mRNA sequence.

ACCESSION AM249090

VERSION AM249090

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Other_ESTs: 2820759, 3prime
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DPF CDNA Library Preparation: Ling
 Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E.
 Consortium (LNLN) DNA Sequencing by: Berkeley MGC sequencing
 project Clone distribution: MGC clone distribution information can

be found through the I.M.A.G.E. Consortium/LLNL at:
www.llnl.gov/bbrp/image/image.html Base Calling / Quality
 Scores: PHRED from University of Washington Genome Center. Vector
 Trimming: cross_match from University of Washington Genome Center
 PHRAP suite. Poly-T identification: palmatch.pl from Berkeley
 Drosophila genome project. University of Washington Genome Center:
<http://www.genome.washington.edu>
 Plate: LCM5 row: A column: 16
 High quality sequence stop: 455.

FEATURES

source

Location/Qualifiers
 1..480
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="2820759"
 /clone_id="NH_MGC-7"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Lung; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GCCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT

68 a 150 c 154 g 108 t

ORIGIN

Query Match 10.6%; Score 30; DB 10; Length 480;
 Best Local Similarity 100.0%; Pred. No. 0.00017;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 88 cccatccctgcctccatctcagagatc 117
 ||||||||||||||||||||||||||||
 Db 451 CCCATCCCTGCTCCATCTCAGAGATC 480

RESULT 13

BG672321

LOCUS BG672321 538 bp mRNA EST 30-APR-2001
 DEFINITION DRNCFD10 Rat DRG Library Rattus norvegicus cDNA clone DRNCFD10 5',
 mRNA sequence.
 ACCESSION BG672321
 VERSION BG672321.1 GI:13894420
 KEYWORDS EST.

SOURCE

ORGANISM

Norway rat.
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 538)
 Xiao,H.S., Han,Z.G., Zhang,F.X., Huang,Q.H., Lu,Y.J., Bao,L., Fu,G.,
 Guo,C., Yan,Q., Jin,S.X., Zhu,Z.D., Xu,X.R., Li,N.G., Chen,Z., and
 Zhang,X.

TITLE

Distinct gene expression profiles of rat dorsal root ganglion
 induced by peripheral nerve axotomy

JOURNAL

COMMENT

Unpublished (2001)
 Contact: Zhang Xu
 Laboratory of Sensory System
 Institute of Neuroscience
 320 Yue Yang Road, Shanghai 200031, P.R.China
 Tel: 86-21-64748700-121
 Fax: 86-21-64713446

Email: xu.zhang@ion.ac.cn

This clone is also available at Chinese National Human Genome
 Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park,
 Pudong New Area, P.R.China. Please contact with Zhang Xu
 (xu.zhang@ion.ac.cn) or Han Zeguang (hanzegu@sh.sh.cn)

PCR Primers
 FORWARD: T3
 BACKWARD: T7

Seq primer: T3
 POLYA-No.

FEATURES

source

Location/Qualifiers
 1..538
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone_image="DRNCFD10"
 /clone_id="Rat DRG Library"
 /sex="male"
 /tissue_type="dorsal root ganglion"
 /dev_stage="adult"

BASE COUNT

118 a 146 c 138 g 136 t

ORIGIN

Query Match 10.3%; Score 29; DB 11; Length 538;
 Best Local Similarity 100.0%; Pred. No. 0.00054;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 189 gaccacaagatttaccagcatcagaga 217
 ||||||||||||||||||||||||||||
 Db 109 GACCAAGAGTTTACACATCATGAGACA 137

RESULT 14

BE231636

LOCUS BE231636 546 bp mRNA EST 10-JUL-2000
 DEFINITION 136437 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION BE231636
 VERSION BE231636.1 GI:9016354
 KEYWORDS EST.

SOURCE

ORGANISM

Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 1 (bases 1 to 546)
 Fahrenheit,K.S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
 Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
 and Keeler,J.W.

REFERENCE

AUTHORS

TITLE

JOURNAL

Design and use of two pooled tissue normalized cDNA libraries for
 EST discovery in swine
 Unpublished (2000)
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390

COMMENT

Email: smtlh@mail.marc.usda.gov
 Single pass sequencing. Bases called and alt. trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -mnscore 18
 and -mismatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCACAGTCACGAG
 Plate: 70 row: A column: 21
 Seq primer: ATTAGGAGACATATAG.

FEATURES

source

Location/Qualifiers
 1..546
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_image="MARC 1P1G"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 library made from pooled tissue from day 11, 13, 15, 20,
 and 30 embryos."

BASE COUNT

76 a 182 c 162 g 124 t 2 others

ORIGIN

Query Match 9.6%; Score 27; DB 10; Length 546;
 Best Local Similarity 100.0%; Pred. No. 0.0054;

Matches 27: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 145 accggctctgtctctaccactgg 171
 |||
 Db 146 ACCGGGCTCTGTCTCTCCTCACCACTGG 172

RESULT 15

AL565390/c

752 bp mRNA

EST 16-FEB-2001

LOCUS

AL565390 LTI_FL013_Fbrn1 Homo sapiens cDNA clone CS0DF005YE02 3

DEFINITION

prime, mRNA sequence.

ACCESSION

AL565390

VERSION

AL565390.1 GI:12916718

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

JOURNAL

Full-length cDNA libraries and normalization

COMMENT

Unpublished (2001)

CONTACT

Genoscope - Centre National de Sequencage

Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

FEATURES

1..752

Source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DF005YE02"

/clone_lib="LTI_FL013_Fbrn1"

/dev_stage="pooled tissue from post conception fetuses (20

week, 24 week and 26 week)"

/lab_host="DH10B"

/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand

cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-stranded cDNA was digested with Not I

and cloned into the Not I and Eco RV sites of the

pCMVSPORT 6 vector. Library was constructed by Life

Technologies. Contact : Feng Liang Life Technologies, a

division of Invitrogen 9800 Medical Center Drive Rockville

, Maryland 20850, USA Fax : (1) 301 610 8371 Email :

fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT

177 a 206 c 242 g 118 t 9 others

ORIGIN

177 a 206 c 242 g 118 t 9 others

Query Match

9.2%; Score 26; DB 10; Length 752;

Best Local Similarity 100.0%; Pred. No. 0.017;

Matches 26: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 97 tggctctcatgtcagagatctcc 122

|||||

Db 615 TGGCTCTCATGTGTCAGAGATCTTCC 590

Search completed: February 13, 2002, 18:48:39
 Job time: 18651 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 12:47:53 : Search time 4552.73 Seconds
(without alignments)
6569.550 Million cell updates/sec

Title: US-09-516-493-6

Perfect score: 1813
Sequence: 1 aactlgtgcgcgcgcgtct.....aaaaaaaaaaaaatlcc 1813

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hvg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_om: *
20: em_or: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vl: *
30: em_hvg_hum: *
31: em_hvg_inv: *
32: em_hvg_rnd: *
33: em_hvg_hum: *
34: em_hvg_inv: *
35: em_hvg_rnd: *
36: em_hvg_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

8

SUMMARIES

Result No.	Score	Match	length	DB	ID	Description
1	1770	97.6	1873	9	HS245937	AJ245937 Homo sapi
2	1770	97.6	2217	6	AX076667	AX076667 Sequence
3	1757.8	97.0	2080	6	AX179740	AX179740 Sequence
4	1397.2	77.1	1508	9	HS217801	Y17801 Homo sapien
5	1033.8	57.0	2087	6	AX076669	AX076669 Sequence
6	1033.8	57.0	2087	10	RN0245935	AJ245935 Rattus no
7	1027.4	56.7	1490	10	MMU17802	Y17802 Mus musculu
8	1027.4	56.7	1843	10	AF232061	AF232061 Mus muscu
9	1024.2	56.5	2072	6	AX076671	AX076671 Sequence
10	1024.2	56.5	2072	10	MMU245936	AJ245936 Mus muscu
11	1020.8	56.3	2189	10	AB033418	AB033418 Rattus no
12	855.4	47.2	1461	6	AX191507	AX191507 Sequence
13	782.8	43.2	1012	4	AF321324	AF321324 Bos tauru
14	589	32.5	789	6	AX191497	AX191497 Sequence
15	545.6	30.1	22370	9	AL445222	AL445222 Human DNA
16	346.4	19.1	1541	6	AX076677	AX076677 Sequence
17	343	18.9	1580	9	HS217803	Y17803 Homo sapien
18	343	18.9	2487	9	HS2011372	AJ011372 Homo sapi
19	327	18.0	2011	6	AX076679	AX076679 Sequence
20	128.6	7.1	127914	2	AC093093	AC093093 Oryza sat
21	128.6	7.1	147706	8	AC083945	AC083945 Oryza sat
22	123.2	6.8	378	6	AX072418	AX072418 Sequence
23	117.4	6.5	120	11	G20347	G20347 human STS A
24	115.8	6.4	150613	2	AP003878	AP003878 Oryza sat
25	112.2	6.2	175506	3	AC008004	AC008004 Drosophill
26	111.4	6.1	1357	8	AB052885	AB052885 Oryza sat
27	108.4	6.0	154180	8	AP000399	AP000399 Oryza sat
28	108	6.0	42845	2	AC017247	AC017247 Drosophill
29	108	6.0	297385	3	AE003540	AE003540 Drosophill
30	107.2	5.9	1463	1	KOARAE	X79598 K. oxyloca a
31	106.6	5.9	132092	2	AP004011	AP004011 Oryza sat
32	105	5.8	177816	2	AC017643	AC017643 Drosophill
33	101.4	5.6	2481	8	CKH0P1	Y07520 Chlorella k
34	96.6	5.3	185413	2	AC010025	AC010025 Drosophill
35	91.2	5.0	2230	4	SSGLUPP	X17058 Sus scrofa
36	90.2	5.0	42415	9	AC001644	AC001644 Genomic s
37	90.2	5.0	44762	9	AC002325	AC002325 Homo sapi
38	90	5.0	2856	9	HUMGLUTRN	K01395 Human (HepG
39	90	5.0	2893	6	AX098209	AX098209 Sequence
40	88	4.9	89787	2	AC020266	AC020266 Drosophill
41	88	4.9	170869	3	AC011696	AC011696 Drosophill
42	88	4.9	171831	3	AC007473	AC007473 Drosophill
43	88	4.9	278196	3	AE003825	AE003825 Drosophill
44	87.2	4.8	10793	1	AE005758	AE005758 Caulobact
45	87	4.8	63855	2	AC016794	AC016794 Drosophill

ALIGNMENTS

RESULT 1
LOCUS HSA245937 1873 bp mRNA 18-FEB-2000
DEFINITION Homo sapiens mRNA for glucose transporter (GLUTX1 gene).
ACCESSION AJ245937
VERSION AJ245937.1 GI:7018305
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1873)
AUTHORS Ibberson,M., Uldry,M., and Thorens,B.
TITLE GLUTX1, a novel mammalian glucose transporter expressed in the
central nervous system and insulin-sensitive tissues
J. Biol. Chem. 275 (7), 4607-4612 (2000)
MEDLINE 20138191
JOURNAL 2 (bases 1 to 1873)
REFERENCE Ibberson,M.R.
AUTHORS Direct Submission
TITLE

Db 1748 TGAAGTCTTTGGAGGTGGTGGTGGCAATTCGGTCTCTGACCGGCTGCGCT 1807
Oy 1743 TATCGGAGGAGAAATTTGTTGCAATAAAGAGCTGACAGAGAAATCAAAAAA 1802
Db 1808 TATCGGAGGAGAAATTTGTTGCAATAAAGAC-TGACACAGAAATCAAAAAA 1866
Oy 1803 aaaaaa 1809
Db 1867 AAAAAA 1873
RESULT 2
AX076667 2217 bp DNA PAT 06-FEB-2001
LOCUS Sequence 1 from Patent WO0104145.
DEFINITION AX076667
ACCESSION AX076667.1 GI:12711198
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2217)
AUTHORS Thorens,B., Ibberson,M. and Uldry,M.
TITLE Glutx polypeptide family and nucleic acids encoding same
JOURNAL Patent: WO 0104145-A 1 18-JAN-2001;
University of Lausanne (CH)
FEATURES
source location/Qualifiers
1..2217
/organism="Homo sapiens"
/db_xref="taxon:9606"
348..1781
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28495.1"
/db_xref="GI:12711198"
/translation="MTPDEPETQPLLGPGSGAPKGRNVLAAFAALGFLSFGFAL
GYSPPAIPSLQRAAPARLDDAASWFGAAVATGAAAGVLGWLVDRAGRKLSLL
CSVPFAGFAVITAAQDVMMILGRLTLGLAGVASLVAPEY;SEIAYPAVGLGSG
VOLAVVIGILLAYLAVGLEMRMLAVILGCVSPSLIMLCMPETPRLLQHRQEA
MAALRPLMGSEQWEDPPIGAEOSPHLALLRQPTLYKPIIIGVILMFQOLSGVNAV
FYATTFIEBAKFKRSSLASVYVGIQVLFYVAALINDRAGRLLLVLSGVYKFSIS
AFGAYEFKLTGGGPNSSSHVAISAPVSAOPDASVGLMVAAGNCLTLAGAVAGMPI
PMLMSEIFPLHNVGAVTGICVLNMLMFLVTKEFSLMEVLRPYGAFWLASFCIF
SVLFTLFCVEIKKTLLEQITAFHEGR"
BASE COUNT 320 a 750 c 672 g 475 t
ORIGIN
Query Match 97.6%; Score 1770; DB 6; Length 2217;
Best Local Similarity 99.4%; Pred. No. 4.2e-269;
Matches 1797; Conservative 0; Mismatches 6; Indels 4; Gaps 2;

Oy 306 Tgctgtcgtgaggccgcctcctcaaccgctctgacctgctgctcctccctagctgccc 365
Db 712 Tgctctgctggggggccgcttcctcaaccgctctgacctgctgctcctccctagctgccc 771
Oy 366 cagctacatctcgaatctgacctaccagcagctcgggggttctgctcctctgtgtgc 425
Db 772 CGGTTCATCTCCGAAATCGCTACCCAGCAGCTCCGGGGGTTGCTCCGCTCTGTCG 831
Oy 426 agctaatgtctgctgagcctcctcctgacctaccctgagcagcgtgggtgctgagc 485
Db 832 AGCTAATGTCGTCTGGCATCTCTGCGCTACCTTGCAAGCTGGGTCTGAGAGTGGC 891
Oy 486 gctgagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 545
Db 892 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 951
Oy 546 Tgcccgaagaccgcgcctctctgctgctgctgctgctgctgctgctgctgctgct 602
Db 952 Tgcccgaagaccgcgcctctctgctgctgctgctgctgctgctgctgctgctgct 1011
Oy 603 Tgctgctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 662
Db 1012 Tgctgctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1071
Oy 663 gcttcaacctgacctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 722
Db 1072 GCTTTCACCTGGCCCTGCTGGCGGAGCCCGGATCTACAGCCCTTTCATCATGGCGCT 1131
Oy 723 cctctgacctgacctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 782
Db 1132 CCTCATGCGCTTCAGAGAGCTGCTGGGGGTCACAGCCCTGATGTTTCATGAGAGACCA 1191
Oy 783 tcttgaagagccagcttcaaggaacagcagcctgctgctgctgctgctgctgctgct 842
Db 1192 TCTTTGAAGAGGCCAAGTTCAAGAGACAGCAGCTGCTGCTGCTGCTGCTGCTGCT 1251
Oy 843 aggtgctgtcaacagctgtgctgctgctgctgctgctgctgctgctgctgctgctgct 902
Db 1252 AGGTCTGTTCACAGCTGTGGCGGCTCTCATCATGACAGACAGCGGGGAGGCTGCTCC 1311
Oy 903 Tgctctgtcaggt 962
Db 1312 Tgctgttgcaggt 1371
Oy 963 Tgaccagaggtgacctgtgacctgctgcaagctgctgcaagctgctgctgctgctgct 1022
Db 1372 TGACCCAGAGTGCCCTGGAACATCTCGACAGTGGCCATCTCGGCGCTGCTGTGCAC 1431
Oy 1023 agcctgtgtatgcaagctgaggtgctgctgctgctgctgctgctgctgctgctgct 1082
Db 1432 AGCTGTGTGATGCCAGCTGGGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCA 1491
Oy 1083 Tgctgctgt 1142
Db 1492 Tgctgctgt 1551
Oy 1143 cctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1202
Db 1552 CTCTGATCTCAAGGCGGTGGGAGCAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 1611
Oy 1203 Tctctgtaacaaagagctcaagcagctcaatgagagctcctcaagcagctcaatgagc 1262
Db 1612 TTTCTGTGACCAAGAGATTACAGCAGCTCATGAGTCTCTGAGGCCCTATGAGGCTTCT 1671
Oy 1263 ggtctgctcctgcttctgcatctcagtgctgtcttcaacttgctgtgtgtgtgtgtgt 1322
Db 1672 GCTTGTGCTCGCTTTCATCTTCAGTGTCTTTTACATTTGTTGTGTGTGCTGTGAAA 1731
Oy 1323 ctaaggaagagctctggaacaaatcaacagccatcttgaagggcgatgaacagccact 1382
Db 1732 TTTAAAGGAAGACCTCTGGAACAATCAGACGCCATTTTGAAGGGGATGAGACGCCACTCA 1791
Oy 1383 ctaaggagatgagcaagcctgtgactcaagctgaggcccaagcccaagccctgtgctgc 1442

|||||
Db 1792 TTAGGGGATGAGCAAGCCCTGTACTTCAAGTTTGGGCCCAAGCCCAAGCCCTGCTGC 1851
QY 1443 cccagggagccagaaatccagcccttgagccttgctgagggctccctctctgt 1502
Db 1852 CCCAGGGAGACCGAATTCAGCCCTTTGGAGCTTGGTGTGCAAGGGTCCCTCTCTGT 1911
QY 1503 catgctccctccagcccaatgagcccgaggtcagagctcaactgctcctgtccagctcc 1562
Db 1912 CATGCTCCCTCCAGCCATGACCCGGGGCTAGAGGCTCTCACTCTCTCTCTCAGCTCC 1971
QY 1563 tgcgtgctcctgaggaactcagaacacttgagcttgacagctcgagctcagagccctc 1622
Db 1972 TGGTGTCTCTGTAGAGACTCAGAACACCTTTCAGCTTTCAGACACCTGGGGTACGCCCTC 2031
QY 1623 catgcaaggaactaaagcagcgaggaagagtgaggtcctcctcagatcttgctctgagc 1682
Db 2032 CATGCCAGAGACTAAAGCAGCGGAGAGAGAGGTGGGCTCTAGATCTTGTCTTCTGGC 2091
QY 1683 tggaggtcctcttgaggttgaggtgctgagcactcagctcctcctcagcgagctgct 1742
Db 2092 TGGAGTGTCTTGTGAGGTGGTGTGCTGGGCAATTCGCTCTCTCTCAAGCGGCTGCC 2151
QY 1743 taccggaaggaatctgttgcacaaataaagacgtgacacaggaataaaaaa 1802
Db 2152 TATCGGAGAGAAATTTGTTTGCCTCAATTAAGAC-TGACACAGAAATCAAAAAA 2210
QY 1803 aaaaaa 1809
Db 2211 AAAAAA 2217
RESULT 3
AX179740 2080 bp DNA PAT 06-AUG-2001
LOCUS AX179740 Sequence 28 from Patent WO0146258.
DEFINITION AX179740
ACCESSION AX179740
VERSION AX179740.1 GI:15132104
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2080)
AUTHORS Baughn, M.R., Burford, N., Au-Young, J., Lu, D.A., Yang, J., Reddy, R.,
Lai, P., Hillman, J.L., Azimzal, Y., Yue, H., Nguyen, D.B., Yao, M.G.,
Gandhi, A.R., Tang, Y.T. and Khan, F.A.
TITLE Transporters and ion channels
JOURNAL Patent: WO 0146258-A 28-28-JUN-2001;
Incyte Genomics, Inc. (US)
FEATURES
Source Location/Qualifiers
1..2080
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 1416107CB1"
BASE COUNT 300 a 688 c 632 g 460 t
ORIGIN
Query Match 97.0%; score 1757.8; DB 6; Length 2080;
Best Local Similarity 99.6%; Pred.No. 3,6e-267;
Matches 1783; Conservative 0; Mismatches 3; Indels 4; Gaps 2;
QY 6 gggcgccgagctctctcctcgccgcttcgagccttgagccttgagcctcagctcagctc 65
Db 110 GCGGCGCCGCGCTTCTCTCGCGGCTTCGCGGCTTCGCGGCTTCGCGGCTTCGCGGCTTC 169
QY 66 tggagctgagctcagcctcccgagcactcctagctcagcgagcgagcccgagcc 125
Db 170 TCGGCTGCTGCTACAGCTCCCGGCAATCCCTAGCTCAGGCGCCCGCGCGCGCC 229
QY 126 cggcgctgagcagcgagcctcctctgttcgaggtgctgagccttgagctgagcgag 185
|||||

Db 230 CGGCGCTGAGCAGACGCCCGGCTCTGTTGGGGCTGTCTGACCCCTGGTGGCCGCG 289
QY 186 cggggagagagctcggagcgtctgagctgagcagcgagcgagcagcagcagcctctg 245
Db 290 CCGGGGAGAGCTGGGGCGCTGGCTGGTGGACCGCGCGCGCGCGCGCGCGCGCGCGCG 349
QY 246 tgtgctcgttgccctctgagccagcttgagcagctcagcagcgagcagcagcagcagc 305
Db 350 TGTGCTCCGCGCTTCTGCTGGCGGCTTTCCTCATACCGCGCGCGCGCGCGCGCGCG 409
QY 306 tgtctgctgagggagcgagcctcctcagcgagcctgagcctgagctgagctgagcctc 365
Db 410 TGGTGTGGGGGGCGCGCTCTCTACCGCGCTGGCGCTGGCTGGCTGGCTGGCTGGCG 469
QY 366 cagctcagcactcctcagaaatccagcagcagcagcagcagcagcagcagcagcagc 425
Db 470 CGGTTTACATCTCCGAATTCGCTTACCCAGCACTCCGGGGGTTGCTCGGCTCTGTTG 529
QY 426 agctaaatgctgctgagcagcactcctcagcagcagcagcagcagcagcagcagcagc 485
Db 530 AGCTAATGCTGCTGCTGCGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 589
QY 486 gctgagctgagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 545
Db 590 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 649
QY 546 tggcagagagcccgagcctcctcagcagcagcagcagcagcagcagcagcagcagc 602
Db 650 TGGCGGAGACCGCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 709
QY 603 tggagctcgttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 662
Db 710 TGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 769
QY 663 gcttcaactgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 722
Db 770 GCTTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 829
QY 723 cctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 782
Db 830 CCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 889
QY 783 tcttgaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 842
Db 890 TCTTTGAAGAGGCAAGTTCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 949
QY 843 aggtctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 902
Db 950 AGGTGCTGTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1009
QY 903 tggctctgagcagctgagcagcagcagcagcagcagcagcagcagcagcagcagc 962
Db 1010 TGGTCTTGTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1069
QY 963 tgaaccagggatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1022
Db 1070 TGACCCAGGCTGGCCCTGCAACTCTCTGCACTGCTGCACTGCTGCACTGCTGCACT 1129
QY 1023 agcctgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1082
Db 1130 AGCTGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1189
QY 1083 tggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1142
Db 1190 TCGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1249
QY 1143 ctctgagctcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1202
Db 1250 CTCTGATGTCAGAGGCTGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1309
QY 1203 tctctgagcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1262
Db 1310 TTTCTGTGACCAAGAGTTTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1369
|||||

Db	870	GCATTTTGAGAGAGGCCAAGTTTCAAGACAGACAGCCTGGCTCGTACTGTGGGCA	TTAA	929
Qy	840	lccagcgtcgtlccacagcgtlbgcggctcctccatccatgacagacagcagcagcagcgtc		899
Db	930	TCCAGGTCCTTTCACGTCTGCTGGCGGCCCTCATTCATGACGACGAGCGCGAAGGCTGC		989
Qy	900	lccgtgtctctcaggt		959
Db	990	TCTGGCCTTCTGGCGGT		1049
Qy	960	agctgaccacagggtgagccttgacacccctccgacgctgtgtgtgtgtgtgtgtgtgtgt		1019
Db	1050	AACGTACCCACAGGCTCCCCACGACACTCTCCACAGTAGGCTT--GGTCCCATCCGCG		1106
Qy	1020	cacagcctgtgtatgacacagcgtgagggctgagctgagctgagcgtgagcgtgtgtgtgt		1079
Db	1107	CGGACCTGTGGATGTCCAACTGGGAGACTGGCTGTGGCTGTATGAGTACATGTGCTCTT		1166
Qy	1080	lcatgcgcgctcttgaggtgagctgagggcccatcccttgctcctcatgtatcagatct		1139
Db	1167	TGATGCGTGGCTTTGGCGGTGGCTGGGAGCCATCCCTGGCTGTATGACATGTGCTCTT		1226
Qy	1140	lccctctgcatgtcaagggcgttgagcagcagcgtcgtgtgtgtgtgtgtgtgtgtgtgt		1199
Db	1227	TGCCCTGTCATGTCAAGGCTGTGCTACCGGCATCTGTGTCTTCCACCACTGCTTCATGG		1286
Qy	1200	cctttctgctgacccaagaggttgaagagccctcatgagaggtgtgtgtgtgtgtgtgtgt		1259
Db	1287	CCTTTTCTAGTACCAACAAAGATTCCAGCGACGCTCATGTGATGATCTCAGACCTTACGGTCC		1346
Qy	1260	lctggtctgctccctccgcttctcgtcatcttccagtgctccttccatctgtctgtgtgt		1319
Db	1347	TCTGGCTACCGGCTGGCTTTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT		1406
Qy	1320	aaactcaaggaagacgtctggaacaatcacagcccatlctgaaggagcgtgacagccac		1379
Db	1407	AGACTAAAGCGACGACTCTGTGACAACTCAGACGCCATTTCGAGGGAGATGACAGACCC		1466
Qy	1380	lcaactgaggagatgagc		1396
Db	1467	TTTTCTGTGCTGGGAGC		1483
RESULT	8			
AF232061				
LOCUS	AF232061	1843 bp	mRNA	23-JUN-2000
DEFINITION	Mus musculus glucose transporter GLUT8 mRNA, complete cds.			
ACCESSION	AF232061			
VERSION	AF232061.1	GI:8671757		
KEYWORDS				
SOURCE				
ORGANISM				
		house mouse.		
		Mus musculus		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;		
		1 (bases 1 to 1843)		
REFERENCE				
AUTHORS		Carayannopoulos,M.O., Chi,M.M., Cui,Y., Pingsterhaus,J.M.,		
		McKnight,R.A., Mueckler,M., Devaskar,S.U. and Moley,K.H.		
TITLE		GLUT8 is a glucose transporter responsible for insulin-stimulated		
		glucose uptake in the blastocyst		
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 97 (13), 7313-7318 (2000)		
MEDLINE		20319023		
REFERENCE		2 (bases 1 to 1843)		
AUTHORS		Moley,K.H., Carayannopoulos,M.O. and Cui,Y.		
TITLE		Direct Submission		
JOURNAL		Submitted (07-FEB-2000) OB/Gyn, Washington University, 4566 Scott		
		Ave, St. Louis, MO 63110, USA		
FEATURES				
source		1. 1843		
		/organism="Mus musculus"		
		/strain="129"		
		/db_xref="taxon:10090"		
		1. 48		

[illegible]

[illegible]

[illegible]

RESULT	14				
LOCUS	AX191497				
DEFINITION	AX191497	789 bp	DNA	PAT	15-AUG-2001
ACCESSION	AX191497	Sequence 19 from Patent WO0149728.			
VERSION	AX191497.1	GI:15209679			
KEYWORDS					
SOURCE		human.			
ORGANISM		Homo sapiens			
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
		Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
REFERENCE		1 (bases 1 to 789)			
AUTHORS		Kato, S. and Kimura, T.			
TITLE		Human proteins having hydrophobic domains and dias encoding these			
		proteins			
JOURNAL		Patent: WO 0149728-A 19 12-JUL-2001;			
		Protegene Inc. (JP); SAGAMI CHEMICAL RESEARCH CENTER (JP)			
FEATURES					
Source		Location/Qualifiers			
		1..789			
		/organism="Homo sapiens"			
		/db_xref="taxon:9606"			
BASE COUNT		76 a 296 c 259 g 158 t			
ORIGIN					

Query Match	32.5%	Score 589	DB 6	Length 789
Best Local Similarly	97.6%	Pred. No. 2.5e-83		
Matches 598; Conservative	0	Mismatches 15	Indels 0	Gaps 0

[illegible][illegible]

RESULT	15
LOCUS	AL445222
DEFINITION	AL445222 225370 bp DNA PRI 24-Apr-2001 Human sequence from clone RP11-356B19 on chromosome 9, complete sequence.
ACCESSION	AL445222
VERSION	AL445222.9, GI:13810082
KEYWORDS	HTG.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 225370)
AUTHORS	Laird,G.
TITLE	Direct Submission
JOURNAL	Submitted (24-Apr-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Requests: clonerequests@sanger.ac.uk On Apr 26, 2001 this sequence version replaced gi:13277497.
COMMENT	

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TRMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr9>

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 12:47:53 ; Search time 283.77 Seconds

(without alignments)
5477.431 Million cell updates/sec

Title: US-09-516-493-6

Perfect score: 1813
Sequence: 1 aacttcgagccgcgcgtctt.....aaaaaaaaaaaaaatcc 1813

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_1101.*
1: /SID2/gcgcdata/geneseq/geneseqn/NA1980.DAT.*
2: /SID2/gcgcdata/geneseq/geneseqn/NA1981.DAT.*
3: /SID2/gcgcdata/geneseq/geneseqn/NA1982.DAT.*
4: /SID2/gcgcdata/geneseq/geneseqn/NA1983.DAT.*
5: /SID2/gcgcdata/geneseq/geneseqn/NA1984.DAT.*
6: /SID2/gcgcdata/geneseq/geneseqn/NA1985.DAT.*
7: /SID2/gcgcdata/geneseq/geneseqn/NA1986.DAT.*
8: /SID2/gcgcdata/geneseq/geneseqn/NA1987.DAT.*
9: /SID2/gcgcdata/geneseq/geneseqn/NA1988.DAT.*
10: /SID2/gcgcdata/geneseq/geneseqn/NA1989.DAT.*
11: /SID2/gcgcdata/geneseq/geneseqn/NA1990.DAT.*
12: /SID2/gcgcdata/geneseq/geneseqn/NA1991.DAT.*
13: /SID2/gcgcdata/geneseq/geneseqn/NA1992.DAT.*
14: /SID2/gcgcdata/geneseq/geneseqn/NA1993.DAT.*
15: /SID2/gcgcdata/geneseq/geneseqn/NA1994.DAT.*
16: /SID2/gcgcdata/geneseq/geneseqn/NA1995.DAT.*
17: /SID2/gcgcdata/geneseq/geneseqn/NA1996.DAT.*
18: /SID2/gcgcdata/geneseq/geneseqn/NA1997.DAT.*
19: /SID2/gcgcdata/geneseq/geneseqn/NA1998.DAT.*
20: /SID2/gcgcdata/geneseq/geneseqn/NA1999.DAT.*
21: /SID2/gcgcdata/geneseq/geneseqn/NA2000.DAT.*
22: /SID2/gcgcdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1770	97.6	2217	22	Human GLUTX1 codin
2	1757.8	97.0	2080	22	Human transporter
3	1033.8	57.0	2087	22	Rat GLUTX1 coding
4	1024.2	56.5	2072	22	Murine GLUTX1 codi
5	855.4	47.2	1461	22	Human protein havi
6	346.4	19.1	1541	22	Human GLUTX3 codin
7	327	18.0	2011	22	Rat GLUTX3 coding
8	237.2	13.1	1527	20	Human transport-as
9	150.6	8.3	579	21	Human secreted exp
10	137.4	7.6	1577	21	Human secreted pro
11	125.6	6.9	1675	20	Corn hexose carlie

12	123.8	6.8	823	21	AAA23461
13	123.2	6.8	378	22	AAF67134
14	91.2	5.0	421	22	AAH99663
15	90	5.0	1815	18	AAT66495
16	90	5.0	2856	21	AAC65871
17	90	5.0	2893	22	AAF98714
18	86.8	4.8	1167	20	AAZ32199
19	85.2	4.7	1172	20	AAZ32203
20	81	4.5	1167	22	AAF81355
21	75	4.1	1826	21	AAC42332
22	74.2	4.1	2504	22	AAF55869
23	73.8	4.1	382	22	AAF65978
24	73.6	4.1	2587	12	AAO11148
25	71	3.9	1776	20	AAZ32200
26	69.6	3.8	426	22	AAF63774
27	68.8	3.8	24379	18	AAT93095
28	68.8	3.8	24379	19	AAV25925
29	68.8	3.8	114955	20	AAK53491
30	67.8	3.7	30001	18	AAT61016
31	67.8	3.7	30001	20	AAK05110
32	64	3.5	58857	21	AAK58471
33	63.8	3.5	718	21	AAE12490
34	63.2	3.5	1943	18	AAE66496
35	63.2	3.5	2850	20	AAZ24632
36	63.2	3.5	10732	21	AAI10594
37	62.2	3.4	1395	21	AAK45857
38	62.2	3.4	1683	21	AAK39099
39	62.2	3.4	1685	21	AAK45298
40	62.2	3.4	1741	21	AAK63319
41	62.2	3.4	1743	21	AAK45290
42	61.8	3.4	3957	22	AAK09686
43	61	3.4	1755	21	AAK40459
44	60.2	3.3	2592	22	AAF81396
45	59.6	3.3	1208	15	AAO64204

ALIGNMENTS

RESULT 1
AAF55865
ID AAF55865 standard; cDNA: 2217 BP.
AC AAF55865;
DT 17-APR-2001 (first entry)
XX
XX Human GLUTX1 coding sequence.
DE
XX
XX Human: GLUTX: gene therapy; vaccine; hexose transport modulator;
KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS Homo sapiens.
XX
XX WO200104145-A2.
PN
XX 18-JAN-2001.
PD
XX 14-JUL-2000; 2000WO-IB01042.
PF
XX 14-JUL-1999; 99US-0143907.
PR 27-AUG-1999; 99US-0151140.
PR 23-FEB-2000; 2000US-0184285.
PR 13-JUL-2000; 2000US-0616132.
PA (UYLA-) UNTV LAUSANNE.
XX
XX Thorens B, Ibberson M, Uldry M;
XX WPI, 2001-112615/12.
DR P-PDB: AAB66932.
XX

QY	843	aggtgctcgttccaaagctctgagctctctcaatcaltgaaacagaagagcgagctctcc	902
Db	950	aggtgctcgttccaaagctctgagctctctcaatcaltgaaacagaagagcgagctctcc	100
QY	903	tgtctctgtcaagtgtgtgtcatagtgttctaaagcagaatgtccttcggcctacttcaag	962
Db	1010	tgtctctgtcaaggtgtgtgtcatagtgtgtctcaagcagaatgtccttcggcctacttcaag	106
QY	963	tgaaccagaaggttgagcccttgagcaactctccgaacgaatgagccaltctgagccctgtctgaac	102
Db	1070	tgaaccagaaggttgagcccttgagcaactctccgaacgaatgagccaltctgagccctgtctgaac	112
QY	1023	agacctctgtatcgcacagcgcgtgagggcttgagcccttgagcccgatgagctcttca	108
Db	1130	agacctctgtatcgcacagcgcgtgagggcttgagcccttgagcccgatgagctcttca	118
QY	1083	tggccgagcttctcgatggagctgaggggcccataccctctgctctccatgttcaagatcttcc	114
Db	1190	tggccgagcttctcgatggagctgaggggcccataccctctgctctccatgttcaagatcttcc	124
QY	1143	ctcttcgatctccaaagcgctgagcgagacatctgtgtctctcaacacgtgctcatatgacct	1202
Db	1250	ctcttcgatctccaaagcgctgagcgagacatctgtgtgtctctcaacacgtgctcatatgacct	1309
QY	1203	ttctctgtacaaagagatctcaagcagcctcatgtgaggtctctcaagccctatgtgaaccttct	1262
Db	1310	ttctctgtacaaagagatctcaagcagcctcatgtgaggtctctcaagccctatgtgaaccttct	1369
QY	1263	ggcttgacctccgcttctctgcatcttctaaagtgtcccttttcaacttgtctctgtgaccttgaaa	1322
Db	1370	ggcttgacctccgcttctctgcatcttctaaagtgtcccttttcaacttgtctctgtgaccttgaaa	1429
QY	1323	cttaagagaagaagactctgtgaaacaaatctcaagcccattctgaggggcgtatgaaacacatca	1382
Db	1430	cttaagagaagaagactctgtgaaacaaatctcaagcccattctgaggggcgtatgaaacacatca	1489
QY	1383	ctaagggatgagagaagcctctgttgactctcaagccttgagcccaagcccacagaccccttgctgc	1442
Db	1490	ctaagggatgagagaagcctctgttgactctcaagccttgagcccaagcccacagaccccttgctgc	1549
QY	1443	cccagaaggagacaggaatctcagcccttgagaccttggagccttggctctgacagggctccctctctg	1502
Db	1550	cccagaaggagacaggaatctcagcccttgagaccttggagccttggctctgacagggctccctctctg	1609
QY	1503	catgtctccctcccaagcccatbaaccgggggctgaaggagctcaatgtccctctgcttccagctcca	1562
Db	1610	catgtctccctcccaagcccatbaaccgggggctgaaggagctcaatgtccctctgcttccagctcca	1669
QY	1563	tgtctgtgtctcttgaagaaactcagaacaaccttcgaagtcttgcagagacttgagcttcaagccctc	1622
Db	1670	tgtctgtgtctcttgaagaaactcagaacaaccttcgaagtcttgcagagacttgagcttcaagccctc	1729
QY	1623	catgtgcagaagactlaaagcagcggagaagagggttgagccctctgagatcttctgtctcttgagc	1682
Db	1730	catgtgcagaagactlaaagcagcggagaagagggttgagccctctgagatcttctgtctcttgagc	1789
QY	1683	tggaggtgtcttcttgnaggttgggtgtcctgggcaatctgagctgagcctccctcaagcggtgctgt	1742
Db	1790	tggaggtgtcttcttgnaggttgggtgtcctgggcaatctgagctgagcctccctcaagcggtgctgt	1849
QY	1743	tatcgggaagagaatctgttctgccaactaaagagtgagcaacagaagaatca	1792
Db	1850	tatcgggaagagaatctgttctgccaactaaagagtgagcaacagaagaatca	1898

RESULT	3
AAF55866	
ID	AAF55866 standard; cDNA; 2087 BP.
XX	

DT 17-APR-2001 (first entry)
XX

DE Rat GLUTX1 coding sequence.

KW Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;
KW hexose transport disorder; ischemia; diabetes; hyperglycemia; ss;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS Rattus sp.
PN WO200104145-A2.
PN XX
PD 18-JAN-2001.
XX
PE 14-JUL-2000; 2000WO-IB01042.
XX
PR 14-JUL-1999; 99US-0143907.
PR 27-AUG-1999; 99US-0151140.
PR 23-FEB-2000; 2000US-0184285.
PR 13-JUL-2000; 2000US-0616132.
XX
PA (UyLA-) UNIV LAUSANNE.
XX
PI Thorens B, Ibberson M, Uldry M;
XX WPI: 2001-112615/12.
DR P-PSDB; AAB66933.
XX

Nucleic acids encoding GLUTX glucose transporter proteins, useful in the prevention, diagnosis and treatment of hexose transport disorders, e.g. ischemia and diabetes -

Claim 3; Page 71-73; 124pp: English.

The present invention relates to GLUTX proteins (AF55865-AF55871 and AAB66932-AAB66941). The GLUTX proteins are related to the facilitative glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport function. The GLUTX proteins may be used in the diagnosis, prevention and treatment of hexose transport disorders such as Ischaemia, diabetes, hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a neurodegenerative disease. The present sequence is the coding sequence for rat GLUTX1.

Sequence 2087 BP; 355 A; 673 C; 580 G; 479 T; 0 other;

Query Match 57.0%; Score 1033.8; DB 22; Length 2087;
Best Local Similarity 84.8%; Pred. No. 2,1e-190;
Matches 1184; Conservative 0; Mismatches 207; Indels 6; Gaps 2

DQ 6 gcggcgccgcgagcttcttcctgcgcgccttcgcgcgtcgcgccttgaggccaactcagttgacct 65
Db gaggcgccgcgagcttcttcctgcgcgccttcgcgcgtcgcgccttgaggccaactcagttgacct 94
QY tgcgcgtcgcgtacaaactcccgcgccatccctagcctcagcgcgcgcgcgcgcgcgcgcgcgcgc 125
Db tgcgcgtcgcgtacaaactcccgcgccatccctagcctcagcgcgcgcgcgcgcgcgcgcgcgcgc 154
QY cgcgccttcgc 185
Db tgcgccttcgc 213
QY cgcgccttcgc 245
Db cgcgccttcgc 273
QY cgc 305
Db cgc 333
QY tgcgtcgc 365
Db tgcgtcgc 394
QY cggtctaatctccgaatactgctaccacgaagtcgcgggggttgtctcgcgtctcgtgtgtgc 425
Db cggtctaatctccgaatactgctaccacgaagtcgcgggggttgtctcgcgtctcgtgtgtgc 453

[illegible]

AAFS5867	
ID	AAF55867 standard; cDNA; 2072 BP.
XX	
AC	AAF55867;
XX	
DT	17-APR-2001 (first entry)
XX	
DE	Murine GLUTX1 coding sequence.
XX	
KM	Murine: GLUTX, gene therapy: vaccine; hexose transport modulator; hexose transport disorder; ischaemia; diabetes; hyperlycaemia; ss; hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX	
OS	Mus sp.
XX	
PN	WO200104145-A2.
XX	
PD	18-JAN-2001.
XX	
PF	14-JUL-2000; 2000WO-IB01042.
PR	14-JUL-1999; 99US-0143907.
PR	27-AUG-1999; 99US-0151140.
PR	23-FEB-2000; 2000US-0184285.
XX	13-JUL-2000; 2000US-0616132.
PA	(UYIA-) UNIV LAUSANNE.
XX	
PI	Thorens B, Ibberson M, Uldry M;
XX	
DR	WPI: 2001-112615/12.
DR	P-PSDB: AAB66934.
XX	
PT	Nucleic acids encoding GLUTX glucose transporter proteins, useful in the prevention, diagnosis and treatment of hexose transport disorders, e.g. ischemia and diabetes -
PT	
PS	Claim 3; Page 73-74; 124pp; English.
XX	
CC	The present invention relates to GLUTX proteins (AAF55865-AAF5871 and AAB66932-AAB6941). The GLUTX proteins are related to the facultative glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport function. The GLUTX proteins may be used in the diagnosis, prevention and treatment of hexose transport disorders such as ischaemia, diabetes, hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a neurodegenerative disease. The present sequence is the coding sequence for murine GLUTX1.
CC	
CC	
XX	
SQ	Sequence 2072 BP; 351 A; 673 C; 584 G; 464 T; 0 other:
Query Match	56.5%; Score 1024.2; DB 22; Length 2072;
Best Local Similarity	84.8%; Pred. No. 1.5e-188;
Matches 1185; Conservative	0; Mismatches 203; Indels 9; Gaps 3.
OY	6 gggagcgcgagcttcctccgcgcgttcgcgcgtcgccctgggccacatcgaattcgct 65
Db	
	85 gccgcgcgcgggctctccctcgcttccttcctgcgcgcgtctggaccctccaacttcgct 144
OY	tgcgcctcgcgcacagctcccgcgcacatccctagccttgacgcgcgcgcgcgcgcgc 125
Db	
	145 tgcgcctcgcgcacagctcccgcgcacatccctagccttgacgcgcgcgcgcgcgcgc 204
OY	cgcgcctcgcgcacgcgcgcgcctccctcgttcgcgggcttgcctgacacctggatccgcgg 185
Db	
	205 tgcgcctcgcgcacaagtgcgcctccctcgttcgcgggcgtctgacccccctggcgtlgt 264
OY	cgggggggaatgcttgaggcgagcttgctgggtggagcacgcgcgggcgaagctgaaacctttgc 245
Db	
	265 caaggggacatactggcgagcttgcttcctcgtgacgcgttcaggggcgaagctgaagctcttgc 324
OY	tctgcctcctgccttcgcgcgcgttcgcgccttcgcctcatcacccgcgcgcgcgttgtga 305

[illegible]

QY	1380	tcactagggatgagc	1396
Db	1462	tttctgtgctgagc	1478

RESULT	5
AAD12574	

ID AAD12574 standard; cDNA; 1461 BP.

AC AAD12574;

DT 25-SEP-2001 (first entry)

Human protein having hydrophobic domain encoding cDNA clone HP10784.

KW Human/hydrophobic domain: gene therapy; nutritional supplement;
KW cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial
KW multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;
KW haematopoietic; tissue growth activity; Parkinson's disease; cytostatic;
KW Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;
KW haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
KW contraceptive; antifertility; antiinflammatory; ss.

OS Homo sapiens.

FH	key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

```
FT      /*tag= a
```

```
FT      /product= "Human protein having hydrophobic domain"  
FT      /note= "CDS is specifically is claimed in claim 3"
```

PN WO200149728-A2.

PD 12-JUL-2001.

PF 28-DEC-2000; 2000WO-JP09359

PR 06-JAN-2000; 2000JP-0000585

PR 11-JAN-2000; 2000JP-0002299

PR	03-MAR-2000; 2000JP-0058367
EN	03-FEB-2000; 2000JP-0020802

PA (PROT-) PROTEGENE INC.

FM (SAGA) SAGAMIL CHEM RES CENI
XX

Pt nato 5, mltitud 1,
XX

DR WPL; 2001-418333/4
DR P-PSDB; AAE06579.

Human proteins with hydrophobic domains and the nucleic acids

Alzheimer's and inflammation - preventing or

XX
PS C:\ajm 4. Page 287-289. 563nn. English

CC The present sequence is human protein with hydrophobic domain encoding
XX

CC CDNA clone HP10784. The polynucleotide and polypeptide of the invention
CC may be used in the prevention, diagnosis and treatment of diseases

associated with inappropriate polypeptide expression. The polynucleotides may be used to produce the polypeptide by inserting the nucleic acid

CC into a host cell and culturing the cell to express the protein. The polynucleotides and its complementary sequences may also be used as

CC probes in diagnostic assays and also used in gene therapy. The

CC and in assays to identify modulators of polypeptide expression and

CC supplements, to modulate cytokine and cell proliferation activity, to

CC microbial infections and autoimmune disorders such as multiple sclerosis,

480 **tg**tcgagattgctcccccaggcgttcgtgggctctggggccacaccccaqctcatag 539

DT 17-APR-2001 (first entry)
 VV

AA121731	standard; cDNA; 1527 BP.
AA121731	
AA121731	
02-NOV-1999	(first entry)
Human transport-associated protein-1 (TRANP-1) cDNA.	
Transport-associated protein; TRANP; nuclear pore; nuclear transport;	
vesicle trafficking; cancer; cystic fibrosis; multidrug resistance;	
hypercholesterolaemia; diagnosis; treatment; ss.	
Homo sapiens.	
Key	Location/Qualifiers
CDS	320..1471
FT	/tag= a
FT	/product= "Human transport-associated protein-1"
PN	W09941373-A2.
PD	19-AUG-1999.
PF	05-FEB-1999; 99WO-0502527.
PR	11-FEB-1998; 98US-0021764.
PA	(INCY-) INCYTE PHARM INC.
PI	Au-Young J, Bandman O, Baughn MR, Corley NC, Guegler KJ;
PI	Hillman JL, Lal P, Yue H;
DR	WP1: 1999-508646/42.
DR	P-PSDB: AAY31639.
PT	Human TRANP coding sequences, used to treat transport disorders and
PT	cancer
PS	Claim 7; Page 78; 87pp; English.
XX	This sequence represents human transport-associated protein-1 (TRANP-1)
XX	CDNA. The DNA sequence was first identified in a human colon tissue
XX	CDNA library. The full-length cDNA was derived from a series of
XX	overlapping and/or extended cDNA sequences and is a consensus.
XX	TRANP-1 to 9 (AA31639-Y31647) are a novel group of proteins with
XX	chemical and structural homology that are involved in molecular
XX	transport. Various disorders are associated with defects in the transport
XX	of molecules, either intracellularly or to the extracellular
XX	environment. Examples of such disorders include cystic fibrosis,
XX	multidrug resistance, hypercholesterolaemia and certain forms of diabetes
XX	mellitus. Defective nuclear transport may play a role in cancer. For
XX	example, the BRCA1 protein, associated with familial breast cancer, is
XX	normally imported into the nucleus via nuclear pore complexes, but is
XX	aberrantly located in the cytoplasm in breast cancer cells. In other
XX	cancers, cells can secrete excessive amounts of hormones e.g, cancers of
XX	the adrenal medulla can secrete excessive amounts of adrenaline and
XX	noradrenaline, leading to hypertension. TRANP is expressed in cancer
XX	cells, and transport disorders result from either excessive or
XX	insufficient molecular transport. Anti-TRANP antibodies and nucleic acids
XX	encoding TRANP can be used as diagnostic tools for such disorders. TRANP
XX	antagonists can be used to treat or prevent a cancer associated with
XX	increased TRANP expression. Anti-TRANP antibodies can be used directly
XX	as an antagonist or as a targeting mechanism for drugs. Alternatively,
XX	a TRANP antisense nucleotide can be used to treat cancers. A TRANP
XX	agonist or expression vector may be used to treat a disorder caused by
XX	reduced transport of biologically active molecules.
XX	Sequence 1527 BP; 216 A; 519 C; 494 G; 298 T; 0 other;

Query Match	13.1%;	Score 237.2;	DB 20;	Length 1527;
Best Local Similarity	56.2%;	Pred. No. 5.8e-37;		

Matches 522; Conservative 0; Mismatches 378; Indels 28; Gaps 3;

QY	15	gacgtctccgagccgctctccgagccctccgagccacacacacgtctccgctcg	74
Db	133	gagtgatctccgagccacatccgacgagtgctcgacaaattcacgtcttgatgcctcg	192
QY	75	gctcacagctccccgagccacatccctccagcctccgacgagccgagcccccgcgcgtcg	134
Db	193	tctacacatcccccgtcatccacgacctcgtagagcgctctcttgatctccgacgtcatctga	252
QY	135	acgagccgagccgctccctctgtgtccggagctgctgtagacctgggtgtccgcgcgaggag	194
Db	253	ccaaatccacagagcattccctgtgtcttggtgctacgtctccacccctgagagcagagccgagag-c	311
QY	195	tgcctgggcgcgtcgctcggtcgtagacccgagccggagccagagcctcttcgtcgctcg	254
Db	312	tgaagtccacatgatatccctcaacgacccctccctggccggagagctgagacatcatgattccagcg	371
QY	255	tgcacctctgctgagccgagctcttgacgtctacacacacgagccacgagagctgtagatgctgctg	314
Db	372	tgcacgtccggagccgagctgatactgagctgacatgagccgggtgcgacagccctctgtagctgctg	431
QY	315	ggggagccgctccctcaacggccctggcgtctggagtgatgtctccctacgtgagcccggtctaca	374
Db	432	tccgaaagagacgtctgacgggagctctccgagggagctccacagctctgacatccacggtgtagcg	491
QY	375	tctccgaaatccgacctacacagacagctccggggagctgtgctccgactccctgtgctgagctgtag	434
Db	492	tgtctgtagatattgcctcccccacggagcggtctgggggctctccgggggagacacaccccccatacg	551
QY	435	tccgacgtccgagatccctccctccgacctccagacacgagctggagctgctgtagagctgagctgctg	494
Db	552	cagtgcttcgagatccctgtctccctctctacgacctggcctccctcgctgacggctggcgtgctgctg	611
QY	495	ctgtgtcctggagctcgagtgccccctccctccatcgtcgtctctctcattgctgcttcagtcgagag	554
Db	612	ctgtggagccggagggcgccgtctgtctcatctacgtatccctgtcgtcagctcttcacgtcccaact	671
QY	555	ccccgcgctctccctgtgactgacacacagacagggagccagagagccatccgacccctggctctct	614
Db	672	cgcgcgcgtctccctgtcctccctccggggagcagagacagagagccctcgccgctgcgtccgag	731
QY	615	ggggctccacgagcaggggctctggagaaagaccccacacacgcgcgtctgagacagagcttccact--	672
Db	732	tgcgttgaggacggagagctcgatgctcatcctgggtgaggtctgagacatccagagacaaagctgcgga	791
QY	673	-----ggccctgctgcggcagcccgagccatccatacgaagccctca	710
Db	792	gacagagacagacgagctacgtctgggtgagcagagccaccccccacagctgtgctccgagccatca	851
QY	711	tcatctggtgtctccctctbatagagccctccacagagctgtctgggggtggtgcaacgacgctcatgct	770
Db	852	ccgggtgcctctgctgatagagccctccctccgaaagagctgagacagacacacagccccaacgtgct	911
Db	912	acctgcagctcatctcttcgacagacagccgctgtctctgtgcctcccccagagagcagcgccaca	971
QY	828	tctgtgggtgtcatccacaggtgctgtcttcaacagctgtgggcggtctctcatctagacacagagcag	887
Db	972	tccgttgaggccgggtgcgtcctctgtcgtctgctgataccgcccctcacacacatgagacatccgag	1031
QY	888	gggcgagagctgctccctgctgtctgtag	915
Db	1032	gcgcgacaggtgctgctctctctgctccagg	1059

RESULT	9
AAAA44914/C	
ID	AAAA44914 standard; cDNA; 579 BP
XX	
AC	AAAA44914;
XX	

DT 21-Aug-2000 (first entry)
 DE Human secreted expressed sequence tag SEQ ID NO:1489.
 XX
 XX Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;
 KW expressed sequence tag; EST; probe; chemotactic; proliferative;
 KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
 KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
 KW antiviral; antidiabetic; antiallergic; antineoplastic; antiparkinsonian;
 KW antitumor; osteoprotective; neuroprotective; neurotropic; antiparasitic;
 KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;
 KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
 KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
 KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
 KW central nervous system disorder; Alzheimer's disease; stroke;
 KW Parkinson's disease; Huntington's disease; coagulation disorder;
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
 KW tumour; infection; depression; psoriasis; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200021991-A1.
 PN
 XX
 PD 20-APR-2000.
 PE
 XX 15-OCT-1999; 99WO-US24206.
 XX
 XX 15-OCT-1998; 98US-0104436.
 XX
 XX (GENM) GENETICS INST INC.
 PA
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Bowman MR;
 XX
 XX WPI: 2000-317938/27.
 DR
 PT Isolated polynucleotides, and encoded proteins, comprising secreted
 PT expressed sequence tags (SESTs), useful for treating various disorders
 PT such as autoimmune, infectious, and central nervous system disorders -
 PS
 PS Claim 1; Page 586; 803pp; English.
 CC AAA43426 to AAA45925 represent specifically claimed secreted expressed
 CC sequence tags (SESTs), isolated from human, mouse, chicken and rat
 CC tissue sources. The SESTs can have a range of activities depending on
 CC the tissues they were isolated from. The activities include:
 CC chemotactic; proliferative; immunomodulatory; haematopoietic;
 CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
 CC cytostatic; antibacterial; antifungal; antidiabetic;
 CC antiallergic; vulnerable; antitumor; osteopathic; neuroprotective;
 CC neurotropic; antiparkinsonian; antipsoriatic; cerebroprotective;
 CC anticonvulsant; and antidepressant. The SESTs can be used for gene
 CC therapy and in vaccines. The SESTs are useful as probes for the
 CC identification and isolation of full-length cDNAs and genomic DNA
 CC molecules which correspond to the SESTs. Proteins encoded by the SESTs
 CC are useful in assays for determining biological activity and raising
 CC antibodies. They may be useful for treatment of autoimmune disorders
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
 CC osteoporosis, osteoarthritis, central nervous system disorders
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
 CC disease), tumours, bacterial, fungal or viral infections, depression and
 CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
 CC in the exemplification of the present invention.
 CC
 XX
 SQ Sequence 579 BP: 109 A; 184 C; 180 G; 106 T; 0 other;

Oy	1073	tgctcttcatcgccagcgttctggcggcgagcttggggccatccctcggtctctatgca	1132
Db	366	tgctcttcatcgccagcgttctggcggcgagcttggggccatccctcggtctctatgca	307
Oy	1133	gagatcttccttcgcatgtcgaaggcgcgttgagcaacgcatctgcgtctcaccacatg	1192
Db	306	GAGATCTTCCCTTCGCAATGTCTAAGGGCGTGGCACAAGGCAATCTGCCTCCTCAACACTGG	247
Oy	1193	ctcatggccttctcgtgacccaaggatctcagcagctcatggag	1237
Db	246	CTCATGGCCTTTCGTGACCAAGGAGCAATTCACGACCTCATGTGG	202
RESULT	10		
AAC80569			
ID	AAC80569	standard; cDNA; 1577 BP.	
XX			
AC	AAC80569;		
DT	12-FEB-2001	(first entry)	
XX			
DE	Human secreted protein gene 39 SEQ ID NO:49.		
XX			
KW	Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;		
KW	antiproliferative; cytoskeletal; cardiant; vasotropic; cerebroprotective;		
KW	neoptropic; neuroprotective; antibacterial; virucide; fungicide; cancer;		
KW	ophthalmological; autoimmune disease; hyperproliferative disorder;		
KW	cardiovascular disorder; cerebrovascular disorder; wound healing;		
KW	nervous system disorder; aging; chemotaxis; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200058467-A1.		
PD	05-OCT-2000.		
XX			
PE	22-MAR-2000; 2000MO-US07505.		
XX			
PR	26-MAR-1999; 99US-0126502.		
PR	17-DEC-1999; 99US-0172410.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
PI	Rosen CA, Ruben SM, Komatsoulis G;		
XX			
DR	WPI: 2000-611712/58.		
DR	P-PSDB: AAB45158.		
XX			
PT	Nucleic acid molecules encoding human secreted proteins, used in		
PT	preventing, treating or ameliorating a disorder, e.g. Alzheimer's and		
PT	Parkinson's diseases and cancers -		
PS	Claim 1: Page 371-372; 440pts. English.		
XX			
CC	Polynucleotide sequences AAC80531-C80580 represent cDNA encoding human		
CC	secreted proteins AAB45120-B45169. Sequences AAB45170-B45225 represent		
CC	alternative polypeptides encoded by the genes, and amino acid sequences		
CC	to which they are homologous. The genes and proteins have activities		
CC	dependent on the tissues and cells in which they are expressed. Examples		
CC	of their activities include immunosuppressive; antiarthritic;		
CC	antirheumatic; antiproliferative; cytoskeletal; cardiant; vasotropic;		
CC	cerebroprotective; neoptropic; neuroprotective; antibacterial; virucide;		
CC	fungicide; and ophthalmological. The secreted proteins, polynucleotides,		
CC	antagonists and agonists may be useful in treating, preventing and/or		
CC	diagnosing diseases and disorders such as autoimmune diseases		
CC	e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms		
CC	of the breast or liver, cardiovascular disorders e.g. cardiac arrest,		
CC	cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous		
CC	system disorders e.g. Alzheimer's disease, infections caused by		
CC	bacteria, viruses and fungi and ocular disorders e.g. corneal infection.		
CC	The polypeptides can also be used to aid wound healing and epithelial		
CC	cell proliferation, to prevent skin aging due to sunburn, to maintain		
CC	organs before transplantation, for supporting cell culture of primary		

Db	1018	agcttcaccacgagccgctcgatcccttccttcacagacgtccaaaggatcaacgtatca	10777
QY	816	tgcacctcgatcgctgagtgatcatccaaagctgcgtctcagacgtctgagccgtatca	875
Db	1078	tgcttcacgagcccgagctgctctgacaaagctcggatctca---agaaagaagctcttcca	11340
QY	876	tgcacaaagacaaaggcggagagctgcttcgtctctgctcagtgatgacatggtatca	935
Db	1135	tgctcttcgtcatcaacgggtctcgcaagctctcgccacgctcgctgctcaatcgacg	11940
QY	936	cgaatgccttcggcgccttaccatcaagctcaaccaaaggatggccctgagcaactctgcacg	995
Db	1195	tcaagcagctgcagcagccgacaaagctgcttcctccaaaggcggcggacagatgactgagctc	12540
QY	996	tggacatcttgagcgcctgcgtctctgcacacagcctgttagatgacagc-----	10400
Db	1255	agctcatctgagtgacagctatccacgcgcgaagcttcgggacacagccagccagggagacatcg	13140
QY	1041	---tgggggcctgcgtctgagccggatggggacagatgagctcttcaatcgccggctctggg	10997
Db	1315	ccaaaggctaacgcgcggctcgatgagctctcatctgcacgtccatcgccatcgccgcgtcttcgct	13740
QY	1098	tggagcttgagggcccatctcccttgagctctctcatctcagaaagatctctctctctgacatgag	11570
Db	1375	ggtctgctggggggccctcgaggtctgagctcgatgcgttcgacagatctctccgcctggagatccgct	14340
QY	1158	gcgtgagcaacagcatctgctgcctccaccacaaactgagctctcatctgctctgctgcacaaag	12170
Db	1435	cggcgaggacagacatcaacagctctccgctcacaacatgctcttcaacctcttcgacatcgacgg	14940
QY	1218	agttcagacagctctcatgagaggtccctcagccctatgagagcctctgcgtctgcctccgct	12770
Db	1495	cccttcctaccacatgctctgctcaactcaagtctgcgctctcttcaactcttcgcgcgcgtcg	15540
QY	1278	ctctgcatctcagatgctctcttcaactgtctc	1309
Db	1555	tcgtcatcatgacagctcttcatgcgtctctc	1586

RESULT	12
AAA23461	
ID	AAA23461 standard; cDNA: 823 BP.
XX	
AC	AAA23461;
XX	
DT	19-JUN-2000 (first entry)
XX	
DE	cDNA encoding human secreted protein vq1_1, SEQ ID NO:77.
XX	
KW	Human; secreted protein; cancer; tumour; cardiovascular disorder;
KW	blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;
KW	infection; fungal; bacterial; viral; HIV; allergy; arthritis;
KW	neurodegenerative disease; asthma; contraceptive; ss.
OS	
XX	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
FT	CDS
FT	54..371
FT	/*tag= a
FT	/product= "Human secreted protein vq1_1"
XX	
FN	MO200011015-A1.
XX	
PD	02-MAR-2000.
XX	
PE	24-AUG-1999;
XX	99WO-US19351.
XX	
PR	24-AUG-1998;
XX	98US-0097638.
PR	24-AUG-1998;
XX	98US-0097659.
PR	09-SEP-1998;
XX	98US-0099618.
PR	28-SEP-1998;
XX	98US-0102092.
PR	25-NOV-1998;
XX	98US-0109978.

PR 23-DEC-1998; 9805-0113645.
PR 23-DEC-1998; 9805-0113646.
PR 23-AUG-1999; 9905-0379246.
XX
PA (ALPH-) ALPHAGENE INC.
XX
XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
PI
XX
DR WPI: 2000-224657/19.
DR P-PSDB: AA195019.
XX
XX New secreted or transmembrane proteins and polynucleotides encoding
PT them, useful for treating neurodegenerative disorders, autoimmune
PT diseases and cancer -
XX
PS Claim 86; Page 335; 357pp; English.

CC The invention relates to 40 human secreted proteins (AA9549981-995020),
CC and cDNA sequences encoding them (AA23423-423462). The secreted
CC proteins of the invention include those that are thought to be only
CC partially secreted, i.e., transmembrane proteins. The proteins of the
CC invention may exhibit one or more activities selected from the following
CC cytokine activity; cell proliferation; differentiation; immune
CC modulation; haematopoiesis regulation; tissue growth activity;
CC activin/inhibin activity; chemotactic/chemokinetic activity; haemostatic
CC and thrombolytic activity; anti-inflammatory activity; and tumour
CC inhibition activity. The proteins may be administered to patients as
CC vaccines, and the nucleotides may be used as part of a gene therapy
CC regime. Diseases or conditions that may be treated using the proteins or
CC nucleotides of the invention include autoimmune diseases; genetic
CC disorders; haemophilia; cardiovascular diseases; cancer; bacterial,
CC fungal and viral infections, especially HIV; multiple sclerosis;
CC rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome;
CC insulin dependent diabetes mellitus; and allergic reactions such as
CC asthma and anaphylaxis. They may also be used for treating wounds, burns,
CC ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's
CC disease, Parkinson's disease, Huntington's disease and amyotrophic
CC lateral sclerosis (ALS). Proteins with activin/inhibin activity may
CC additionally be useful as contraceptives. Nucleic acid sequences of the
CC invention may be used in chromosome mapping, and as a source of
CC diagnostic primers and probes. The present sequence represents cDNA
CC encoding one of the 40 proteins of the invention.

Sequence 823 BP; 170 A; 257 C; 233 G; 163 T; 0 other;

Query Match	6.8%	Score 123.8	DB 21	Length 823
Best Local Similarity	63.0%	Pred. No. 3.7e-15		
Matches 191; Conservative	0	Mismatches 112	Indels 0	Gaps 0

[illegible]

RESULT 13

AA67134

AA67134 standard; cDNA: 378 BP.

XX AA67134:

DT 09-APR-2001 (first entry)

DE Novel human polynucleotide, SEQ ID NO: 2890.

KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.

XX Homo sapiens.

OS WO200102568-A2.

PN 11-JAN-2001.

PF 30-JUN-2000; 2000WO-US18374.

PR 02-JUL-1999; 99US-0142310.

PR 02-JUL-1999; 99US-0142311.

XX (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Cirenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LM, Strache-Crain B;

DR WPI: 2001-091805/10.

PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -

PS Claim 9; Page 974; 1046pp; English.

CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.

XX Sequence 378 BP; 45 A; 130 C; 117 G; 86 T; 0 other;

Query Match 6.8%; Score 123.2; DB 22; Length 378;

Best Local Similarity 63.5%; Pred. No. 4.2e-15;

Matches 188; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

OY 1053 gctgagccgtggcagcatgtgctcttcacgcgcggttgcgggtgggtgggagccca 1112

DB 8 ggggtccctgtggcgcacacatccttcacatcgtgtaagcgcgtgggtgggtccca 67

OY 1113 tccctgtcctcctatgacagatccttcctcgtcagatcagaggtggtgcagagca 1172

DB 68 tcaactgctgctcatgtctgagctgcctcctgctgacccgtggcgtggtcctcaaggc 127

OY 1173 tctgctcctcaccacacatgctcctccttcctcgtgaccaaagatcagcagcctca 1232

DB 128 tcttcgtgctgagccagctgctcaccgccttccttcaccaaagctcctcagcagtg 187

OY 1233 tggaggtcctcagggccctatgagaccttcctgctcgtccttcgtcatctcagtg 1292

DB 188 tgaagaccttcggcctccagtgctcttccttccttccttcgtgcgacatcgtctggagcc 247

OY 1293 tcccttcacttctcgtcgtccctgaacctaagaagaactcgtgaacaatc 1348

DB 248 tgggttcacagggctgctgctgtgcccagaccagagagcgttcctgagcagatc 303

RESULT 14

AAH99663

AAH99663 standard; cDNA: 421 BP.

XX AAH99663:

DT 16-OCT-2001 (first entry)

DE Human protein encoding cDNA sequence SEQ ID NO:498.

KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anemia;
KW antiagregant; haemostatic; vulnery; antileuc; osteopathic; eczema;
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder; ss.

OS Homo sapiens.

PN WO200153455-A2.

PD 26-JUL-2001.

PF 22-DEC-2000; 2000WO-US35017.

PR 23-DEC-1999; 99US-0471275.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI: 2001-457603/49.

PS P-PSDB: AAM25722.

PS Claim 1; Page 564; 1217pp; English.

CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAH25963. The proteins can have activities based on the tissues and
CC cells they are expressed in, such as: antinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiagregant; haemostatic; vulnery;
CC antileuc; osteopathic; dermatological; antiallergic; antiasthmatic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine

Thu Feb 14 07:44:40 2002

us-09-516-493-6.rng

Page 16

Search completed: February 13, 2002, 13:42:44
Job time: 3291 sec

[illegible][illegible][illegible]

```

1      RESULT      3
2      US-08-125-468-1/c
3      : Sequence 1, Application US/08125468
4      : Patent No. 5589385
5      :
6      : GENERAL INFORMATION:
7      :
8      : APPLICANT: Ryan, Michael J.
9      : APPLICANT: Lotvin, Jason A.
10     : APPLICANT: Strathy, Nancy
11     : APPLICANT: Fantioli, Susan E.
12     :
13     : TITLE OF INVENTION: Cloning of the biosynthetic pathway for
14     : TITLE OF INVENTION: chlorotetracycline and tetracycline formation and cosmid
15     : TITLE OF INVENTION: useful therein
16     :
17     : NUMBER OF SEQUENCES: 1
18     :
19     : CORRESPONDENCE ADDRESSES:
20     :
21     : ADDRESSEE: American Cyanamid Company
22     :
23     : STREET: One Cyanamid Plaza
24     :
25     : CITY: Wayne
26     :
27     : STATE: New Jersey
28     :
29     : COUNTRY: USA
30     :
31     : ZIP: 07470
32     :
33     : COMPUTER READABLE FORM:
34     :
35     : MEDIUM TYPE: Floppy disk
36     :
37     : COMPUTER: IBM PC compatible
38     :
39     : OPERATING SYSTEM: PC-DOS/MS-DOS
40     :
41     : SOFTWARE: Patent Release #1.0, Version #1.25
42     :
43     : CURRENT APPLICATION DATA:
44     :
45     : APPLICATION NUMBER: US/08/125,468
46     :
47     : FILING DATE: 22-SEP-1993
48     :
49     : CLASSIFICATION: 435
50     :
51     : ATTORNEY/AGENT INFORMATION:
52     :
53     : NAME: Tsevdos, Estelle J
54     :
55     : REGISTRATION NUMBER: 31,145
56     :
57     : REFERENCE/DOCKET NUMBER: 31,255-02
58     :
59     : TELECOMMUNICATION INFORMATION:
60     :
61     : TELEPHONE: (201)831-3241
62     :
63     : TELEFAX: (201)831-3305
64     :
65     : INFORMATION FOR SEQ ID NO: 1:
66     :
67     : SEQUENCE CHARACTERISTICS:
68     :
69     : LENGTH: 30001 base pairs
70     :
71     : TYPE: nucleic acid
72     :
73     : STRANDEDNESS: single
74     :
75     : TOPOLOGY: linear
76     :
77     : MOLECULE TYPE: DNA (genomic)
78     :
79     : US-08-125-468-1

```

Query Match	3.7%;	Score 67.8;	DB 1;	Length 30001;
Best Local Similarity	44.7%;	Pred. No. 3e-05;		
Matches 264;	Conservative	0;	Mismatches 327;	Indels 0;
			Gaps	0;
QY	6	gcgcgcgcgcgcgtctctctccgcgcgccttcgcgcgtgcgccttgcccaactcaagcttcgct	65	
Db	20412	GCCCGCGCGGCTTCACTACCCGCGAGATCAATCAGCGGCTGTCGGGCTCTGTTGGCGG	20353	
QY	66	tgcgcgtcgcgcgtcaacgcgtccccgcgcacatccctaagcctcagcgcgcgcgcgcgcgcgc	125	
Db	20332	TGCTCTCTGCGCGCACTGACACAGAGATGTCCTCCACGCGCGTGGTACGATGCGCGACC	20293	
QY	126	cgcgccttgacgcgcgcgcgcgcctcttcgttcttcgcgcgtcgttcgtgaacctggatgcgcgc	185	

QY	317	gqgcgcctcccaacggccctgagcctgagtggtctgacctagtgcccagctctaacac	376
Db	1362	ggcggcctcccaactaacctctccgcagcccgagctccaactcctcgcagcccgctccac	1422
QY	377	tcggaatcgcctaacccagcagctccgggggtctgctcgctctgtatgcactaagtgctc	436
Db	1422	tgctctctcgagcccgagctccacacctcctgctctccgcctctctgtctctcgccctcttc	1481
QY	437	gtcgtagcaatcctccctgagctaacctcgcagagctctgggtgtctgtagagtgagcctgagctgct	496
Db	1482	ctgcctctgcgccctccgcgcctccctcgtcctctgagccctctgcctctgcgcctccctgctccctgccc	1541
QY	497	gtctctggagctgctgctcccccctccctccatctgctctctcatgtgctctcaatgcccagagacc	556
Db	1542	ctctctgcccctcctgctctcgcgcctccctccctcctcctctcctctcctgctctctgcccctccgcgc	1601
QY	557	ccgcgcctctcctgctctgactcaagcaacagcgagcgagagagccatcgccctgcggtctccctgag	616
Db	1602	ctctctctctgctccctgcgcctccctctgcctctctgctctctgctccctctctgctccctctgctctc	1661
QY	617	ggctccgcagacgggctctggagagacccccccatctggggctcgagcagagctcttaacctggccc	676
Db	1662	ctgcgcctctctgcccctctccctgctctccctcgcgcctctcgtctcctgacccctctctgctctctgccc	1722
QY	677	ctgctgcgcgacgcccgggcatctacaagcccttcatcatcggtgtctctccctgtagatggcttc	736
Db	1722	ctctctgctccctgcgcctccctctgcgcctctccctctcctctcctctgctctctgcccctccgcctc	1781
QY	737	cagcagctgtcgggggtctcaacgcgcctgattgtc	769
Db	1782	ctgcgcctccctgcgcctccctctgcgcctctcctgctc	1814

```

1      RESULT      8
2      US-08-910-647-1/c
3      Sequence 1, Application US/08910647
4      Patent No. 6251433
5      GENERAL INFORMATION:
6      APPLICANT: Zuckermann et al.
7      TITLE OF INVENTION: Compositions and Methods for
8      TITLE OF INVENTION: Polynucleotide Delivery
9      NUMBER OF SEQUENCES: 4
10     CORRESPONDENCE ADDRESS:
11     ADDRESSEE: Chiron Corporation
12     STREET: 4560 Horton Street
13     CITY: Emeryville
14     STATE: California
15     COUNTRY: U.S.A.
16     ZIP: 94608-2916
17     COMPUTER READABLE FORM:
18     MEDIUM TYPE: Floppy disk
19     COMPUTER: IBM PC compatible
20     OPERATING SYSTEM: PC-DOS/MS-DOS
21     SOFTWARE: PatentIn Release #1.0, Version #1.30
22     CURRENT APPLICATION DATA:
23     APPLICATION NUMBER: US/08/910,647
24     FILING DATE:
25     CLASSIFICATION: 514
26     ATTORNEY/AGENT INFORMATION:
27     NAME: Fujita, Sharon M.
28     REGISTRATION NUMBER: 38,459
29     REFERENCE/DOCKET NUMBER: 1218.002
30     TELECOMMUNICATION INFORMATION:
31     TELEPHONE: (510) 923-2706
32     TELEFAX: (510) 655-3542
33     INFORMATION FOR SEQ. ID NO.: 1:
34     SEQUENCE CHARACTERISTICS:
35     LENGTH: 9600 base pairs
36     TYPE: nucleic acid
37     STRANDEDNESS: single
38     TOPOLOGY: linear
39     MOLECULE TYPE: DNA (genomic)
40     US-08-910-647-1

```

[illegible]

RESULT 9
US-07-884-811-15/c
Sequence 15, Application US/07884811
Patent No. 5316921
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
TITLE OF INVENTION: SINGLE-CHAIN HERATOCTYE GROWTH FACTOR VARIANTS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/884, 811
FILING DATE: 19920518
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dieger, Giner R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 755.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881

```

RESULT 10
US-07-885-971-15/c
: Sequence 15, Application US/07885971
: Patent No. 5328837
:
GENERAL INFORMATION:
:
: APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
: TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
:
COMPUTER READABLE FORM:
:
: MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
:
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: palin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/885,971
: FILING DATE: 19920518
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:

```

[illegible]

RESULT 11
 US-08-087-783A-15/c
 : Sequence 15, Application US/08087783A
 : Patent No. 5547856
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R
 : TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
 :
 : NUMBER OF SEQUENCES: 22
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Genentech, Inc.
 : STREET: 460 Point San Bruno Blvd
 : CITY: South San Francisco
 : STATE: California
 : COUNTRY: USA
 : ZIP: 94080
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 3.5 Inch, 1.44 Mb floppy disk
 : COMPUTER: IBM PC compatible

Thu Feb 14 07:44:41 2002

us-09-516-493-6.rni

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 12:47:53 ; Search time 2842.4 Seconds (without adjustments)

(without alignments)
6854.101 Million cell updates/sec

Title: US-09-516-493-6
Perfect score: 1813

Sequence: 1 aactgcygcgcgcgtct.....aaaaaaaatltcc 1813

Scoring table:	IDENTITY_NUC	Gapnext 1 0
	Gapopen 10 0	

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Maximum DB seq length: 20000000000

Post-processing:	Minimum Match 0%
	Maximum Match 100%

Listing first 45 summaries

Database :

```

1:  em_estfun:*
2:  em_esthnm:*
3:  em_estlin:*
4:  em_estlom:*
5:  em_estlpm:*
6:  em_estlpl:*
7:  em_estlth:*
8:  em_estlov:*
9:  em_hlcc:*
10: qb_est1:*
11: qb_est2:*
12: qb_hlcc:*
13: qb_hlcc:*
14: em_gss_fun:*
15: em_gss_hnm:*
16: em_gss_inv:*
17: em_gss_plm:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vtr:*
21: em_gss_other:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	801.4	42.2	932	11	B6479842	B6479842 602527377
2	768.6	44.4	949	11	B6749509	B6749509 602507648
3	746.2	41.2	830	11	B6700749	B6700749 602681614
4	701.6	38.7	991	11	B1334832	B1334832 602998955
5	659	36.3	752	10	AL565390	AL565390 AL565390
6	634.8	35.0	828	11	B6478000	B6478000 602522606
7	634.4	35.0	879	11	B6717136	B6717136 602720404
8	633	34.9	934	11	B6910478	B6910478 601501005
9	631	34.8	777	11	B6717034	B6717034 602689039
10	603.6	33.3	712	11	B6290177	B6290177 602348526
11	599	33.0	838	11	B6281777	B6281777 602405299
12	499.2	27.5	801	11	B114765	B114765 602861314

13	483	26.6	664	11	BE986058
14	482.4	26.6	677	11	BG70107
15	477	26.3	651	10	AM955452
16	476	26.3	621	11	BE986074
17	472	26.0	673	10	A1094911
18	471	26.0	480	10	AM249090
19	457	25.2	457	10	A1971919
20	454.6	25.1	594	11	B1402088
21	440.4	24.3	442	10	AM051497
22	439.6	24.2	460	10	AM055123
23	437.4	24.1	451	10	A1380005
24	428.6	23.6	466	10	A1149830
25	425.4	23.5	441	10	A1338697
26	420.8	23.2	595	11	BG079217
27	420.2	23.2	917	11	BE140667
28	408.4	22.5	417	10	AA436091
29	405.4	22.4	421	10	A1205662
30	405.2	22.3	920	11	BF308306
31	399.6	22.0	487	10	AA293722
32	399.4	22.0	519	11	B1404477
33	397	21.9	921	11	BF311341
34	388	21.4	927	11	BF207033
35	385	21.2	498	11	B1403682
36	382.4	21.1	396	10	A1025047
37	380.4	21.0	389	10	A1041348
38	378.4	20.9	823	11	BF120840
39	372.2	20.5	425	10	A11290245
40	367.2	20.3	623	10	BE282246
41	365.6	20.2	477	11	BF400621
42	365.6	20.2	478	11	BE115293
43	363.6	20.1	458	10	AM491310
44	361	19.9	389	10	A1970024
45	359.4	19.8	374	10	AM300975

ALIGNMENTS

```

RESULT      1
BG479842
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
    BG479842      932 bp      mRNA      EST      21-MAR-2001
    602527373F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4650906 5',
    BG479842
    BG479842.1   GI:13412121
    EST.
    human.
    Homo sapiens
    Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
    1 (bases 1 to 932)
    NIH-MGC http://mgc.ncl.nih.gov/.
    National Institutes of Health, Mammalian Gene Collection (MGC)
    Unpublished (1999)
    Contact: Robert Strausberg, Ph.D.
    Email: cgaps-remail.nih.gov
    Tissue Procurement: ATCC
    cDNA Library Preparation: Ling Hong/Rubin Laboratory
    cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
    DNA Sequencing by: Incyte Genomics, Inc.
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LNL at:
    http://image.llnl.gov
    Plate: LCM1433 row: a column: 19
    High quality sequence stop: 820.
    Location/Qualifiers
        1..932
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4650906"
            /clone_11b="NIH_MGC_21"
            /tissue_type="choriocarcinoma"
            /lab_host="DH10B (phage-resistant)"

```



```

Db 181 CCTCATGTCAGAGATCTTCCCTCTGCAATGTCAGAGGCGTGGGACAGGATCTGGCTCT 240
QY 1183 caccactggtcatalgaccttctcgtgaccagaaggttcagcagcctcatgaggtcct 1242
Db 241 CACCAACTGGGTCATGGCTTTCTGTCGACCAAGAGTTTACAGACCTTCATGAGAGTCT 300
QY 1243 cagcctctatgagacctctgacctgacctcgccttctgacctctcaatgaccttctcaac 1302
Db 301 CAGGCGCTATGAGAGCTTCTGGCTTGGCTCCGCTTTCGACATCTTCACTGCTCTTTTAC 360
QY 1303 ttctgtctgtctccctgaaactaaagaaagacttgaaacaatcacagccacttga 1362
Db 361 TTTGTTCTGTGTGCTTGAAGCTAAAGAGAGACTGTGAACAATTCACAGCCATTGGA 420
QY 1363 gggagcatgagacacacacactaagagatgagacagacctgtactcaagcttgagcca 1422
Db 421 GGGGCGATGACAGCCACTCAGGCGATGAGGAGCAAGCCTGTGACTCCAGCTGGGCCCA 480
QY 1423 agccagagacccctgacctgacctcagagagagagcagaatccagcccttgagacctgtct 1482
Db 481 AGCCAGAGCCCTGCTGCCCCAGGGAGGAGCCAGAAATCCAGCCCTTGGAGCCTTGCT 540
QY 1483 gcagaggtccctctctcgtatgctcctcagaccatgaccgggagctagagctca 1542
Db 541 GCAGGTCCTCTCTCTCTGTCATGCTCCCTCAGCCATGACCCGGGCTAGAGGCTCA 600
QY 1543 ctgctctctgtcagacctgacctgacctgacctgagactcagaaacctctgagaccttg 1602
Db 601 CTGCTCTCTGTTCAGCTCCGCTGCTCTGCTGAGAGACTCAGAGACCTTCGAGCTTTG 660
QY 1603 cagacctgagctcagacctcctcagcagaagactaaagcagcagagagagagctgagccctc 1662
Db 661 CAGACTGCTGCTCAGCCCTCCATGCGCAAGACTAAAGCAGCGAGAGAGAGAGAGAGCTG 720
QY 1663 tagatcttctctctgctgctgagaggtcttcttgaaggttggtg---ctggagcttgg 1719
Db 721 TAGGATCTTCTCTCTCTGCTGAGAGTCTTCTTTCGAGGTTGGGCTGCGCATTTGG 780
QY 1720 tgcct-cctctcagcagcgtcctcctatcagagagaattgcttgccaataaagagct 1778
Db 781 ACGTCCCTCTCAGCGGCTGCTTATCGGAGACCAATTCGTT--GCCAATCAGAGACT 838
QY 1779 gacacagaataca 1792
Db 839 GACACAGAGAGATCA 852

RESULT 3
LOCUS BG700749 830 bp mRNA EST 07-MAY-2001
DEFINITION 602681616F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4814451 5',
ACCESSION BG700749 mRNA sequence.
VERSION BG700749.1 GI:13970402
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo.

REFERENCE 1 (bases 1 to 830)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgs@ncl.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

```

```

FEATURES
source
Plate: LLAM10709 row: p column: 04
High quality sequence stop: 789.
Location/Qualifiers
1..830
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4814451"
/clone_id="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcag
); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.5 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI/NHORI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 104 a 271 c 266 g 189 t
ORIGIN
Query Match 41.2%, Score 746.2; DB 11; Length 830;
Best Local Similarity 97.1%; Pred. No. 1.6e-105;
Matches 803; Conservative 0; Mismatches 18; Indels 6; Gaps 4:
QY 337 ggcctgaggtgtgtcctcctcagtagtgcccggtcacaatccgaaatcgcctaccagc 396
Db 4 GGCCTGGGGGTGTTGGCTCCCTAGTGGGCCCGGCTACATCTCCGAATCGCTACCCAGC 63
QY 397 agtcagaggtgtgtcgcgtcctcgtgtgcaagctaaatgtgtgtgtgtgtgtgtgtgtgt 456
Db 64 AGTCGGGGGTTGCTCGGCTCCTGTGTGACAGTATGCTGCTGCTGCTGCTGCTGCTGCT 123
QY 457 ctacctgcaagcttaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 516
Db 124 CTACCTGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183
QY 517 ctccctatgtcgtctcctcagtagtctcagcagagcccggtctcctcgtcagctca 576
Db 184 CTCCCTCATGCTGCTTCTCATGTGCTTCTCATGCTCCAGACCCCGCTTCTCTGCTACTCA 243
QY 577 gacacagcagcagagagcgt---cgacctgagcttctgtgtgtgtgtgtgtgtgtgtgtgt 633
Db 244 GCACAGGCGCCAGAGGCGATGCGCCGCTGCGCTTCTGTGCGGCTCCGAGACAGGCTG 303
QY 634 ggaagacccccccatcagggctgagcagaagcttcaacctgtgacctgtgcgagcccg 693
Db 304 GGAAAGACCCCCCATCGGGGCTGAGACAGAGCTTTCACCTGGCCCTGCGGACGCCGG 363
QY 694 catctaaagccctcattcattcagtggtctcctcgtgtgtgtgtgtgtgtgtgtgtgtgtgt 753
Db 364 CATCTTACAAGCCCTTCAATCATGAGGCTGCTCCCTGATGAGCTTTCAGCAGCTGTGCGGGGT 423
QY 754 caacgcctcatgtctctatctcagaaagacatcttbaaagagccaaagctcagaagcag 813
Db 424 CAACGCCGTCATGTCTTATGAGAGACCATCTTTGAACAGGCGAGTTTCAGAGACAGCAG 483
QY 814 ccgtgctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 873
Db 484 CCTGCGCTGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
QY 874 catgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 933
Db 544 CATGGACAGAGAGGCGCGAGGCTGCTCTCTGCTTTCAGAGTGTGTCATGAGTGTTCAG 603
QY 934 cagcagtgacctgtgcgtcactcaagctcagcagcagcagcagcagcagcagcagcagcagcag 993
Db 604 CACGAGTCTCTTGGCGGCTTACTTCAACGTAGCCAGAGGTGGCTTGGCACTCTCGCA 663
QY 994 cgtgacatctgcgacctgtctctg--cacagcctg--ttgatgcagcagcagcagcagcagc 1051

```

D6	CGTGGCCATCTCAGGCGGCTGTCTCTTGCCACAGAGCCTGTTCAGAACAGAGCTGGAGCTTGAGCC	723
Oy	1052 tggctt-ggcgctggagcagatgtgctctttcaacgccggccttgtagtgactggaggcc	1110
D6	724 TGGCTGGAGCCGTGGGAGCATGTGCCCTTTCAATCGGCCTTTGGCGGTGGCGGGGAGG	783
Oy	1111 catcccctgacctcatgtcacagatcttcctcctgcagtccaag	1157
D6	784 CATCCCTTGCTCTCATGTCAAGAGATTTCCTCTTGCTGTGCCG	830
RESULT	4	
B1334832		
LOCUS	B1334832	991 bp mRNA EST 30-JUL-2001
DEFINITION	602998959F1 NIH_MGC_12 Homo sapiens CDNA clone IMAGE:5141007 5'	
ACCESSION	B1334832	
VERSION	B1334832.1	GI:15019489
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 991)	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-f@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: Incyte Genomics, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM1347 row: f column: 16 High quality sequence stop: 835. Location/Qualifiers	
FEATURES		
Source	1..991	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:5141007"	
	/clone_1lb="NIH_MGC_12"	
	/tissue_type="cervical carcinoma cell line"	
	/lab_host="DH10B"	
	/note="Organ: cervix; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."	
BASE COUNT	138 a 326 c 311 g 216 t	
ORIGIN		
Query Match	38.7%; Score 701.6; DB 11; Length 991;	
Best Local Similarity	88.8%; Pred. No. 1.1e+98;	
Matches 865; Conservative	0; Mismatches 89; Indels 20; Gaps 5	
Oy	296 gaactgtgatgtcgtcctggaggcgcccttaaccgagcctgcagtgattgacctc	355
D6	1 GGCGTTGGCCGCCGACACTGACGCCCGAGAGACCAGAAGAAACCCACGCCCTTGTGGGCG	60
Oy	356 ctatggccccggtctacatctccgaatcgcctaccacagcagctccggagtgctgcgac	415
D6	61 TCCTGGCGGCGAGGCTACATCTCCGAATGCGCTACCCAGCAAGTCCGGGGTTCGCGGC	120
Oy	416 tcctgttgtcaagataatgtcgtcgtctgcgacatctctcgtgctactgcagagctggatg	475
D6	121 TCCTGTGTCAGCAATAATGTCGTCTGTCGACATCTCTG6CCYACCTG6CACGGCTGGTG	180
Oy	476 ctgaagaaggagctggctgt	535
D6	181 CTGAGATGGCCCTGGCTGGCTGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	240

OY	536	atctgcttcacgcccgaagaccgcgcgtctccctgtgtacatcaagcagcgccaaagagcc	595
Db	241	ATGTGCTTTCATGTCGCCGAGACCCCGCGCTCTGCTGTGACTACAGCACAGGCGCAGAGAGCC	300
OY	596	at- --cgccctgcgggttcctgtgtgagctcgcagcagagcgctgtggaagaccccccatcgga	652
Db	301	ATGGCGCGCCCTGCGGGTTCCTGTGGGGCTCCAGCAGAGGAGCTTGGAAGACCCCCCATCGGG	360
OY	653	gctgagcagaaacttcaacctgagccctgtgtggagagcccgagcatcaaaagccctatc	712
Db	361	GCTGACGAGACTTTCACATCTGGCCCTGTGGGACGCCCGGCATCTTCAAAACCCCTTATC	420
OY	713	atcggatgtccctcgtatgtgaccttcccaagcagctgtgcgggggttcaacgcgcgtcatgttcat	772
Db	421	ATCGGCGCTCTCCGTCATGCGCTTCCAGAGAGCTGTGCGGGGCTCAACGCGCGTCATGTTCTAT	480
OY	773	gcagaagaccatcctttaaagagcgcaagtlcaagagcagcagcctgtgcctgcgttcgtg	832
Db	481	GCAGAGACCACTTTTGAAGAGAGCCAAAGTTCAAGAGACAGACGCCCTGGCTGTGTCGTG	540
OY	883	gggtgtacacaaagtgctgttcaacagctgtgagcgtctcctcaatcagacagagcagagcg	892
Db	541	GGTGTACATCCAGGTCGCTCTTACAGCTGTGGCGGCTCTCATCATGTGACAGACAGCGGG	600
OY	893	aggctgcctccctgctctgtcagtgatgtgtcagtgatgtgtcagcaagatgtccttcgagcc	952
Db	601	AGGCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG	660
OY	953	tacttcaagctgaccccaagaggtgtgcctctgtgcaactctctcgc- acgttgccatctcgcgc	1011
Db	661	TACTTCAAGCTGACGCCAGCGAGGTCGCCCTGTGGCAACCTCCGCAACTGTGCCATCTCGGGCC	720
OY	1012	tgctctctgcac- agcctcctgtatgtacagagcgtgtggg- ctgacctgtgcgcg- tggggcag	1068
Db	721	TGTCTCTGCACAAAGCTGTTCATGTCACAGCTGTGGGGCCTGGCTGTGGCTGTGGGACG	780
OY	1069	catgtgcctctcttcacgcgcgagcttgcgtgtgagct- --gggggccaatccctctgacct	1125
Db	781	CATGTGCTCTTCATGCGCGGCTTTGGGGGTGGGGCTTGGGGGGGCCCATTTCCCTGTGGTCCC	840
OY	1126	catgtcagaagatcttccctctcgtcatgttca- -----aggcggtgtggacaag- -catctg	1176
Db	841	TCATTTGTCAGAGATCTTTCCCTCTGTGATTTGATGTCACCGGGCGTGGCAGCAGGGCATTTCTA	900
OY	1177	gcctctcaacaaacgtgctcatgagccttctcgtatcaagaagatgtcagcagcctcatgga	1236
Db	901	GSTGCCCAACAATGGCTCATGAGCTTTCCCGGGAGC- AGSAGTTTACGCGAGCTCTTGA	959
OY	1237	ggtcctcagagcct 1250	
Db	960	AGTCCCGCAGGCCCT 973	
RESULT	5		
AL565390/c			
LOCUS	ALS65390	752 bp	mRNA
DEFINITION	ALS65390 LTI_F1013_FBRnl Homo sapiens cDNA clone CS0DF005YE02 3		
ACCESSION	AL565390		
VERSION	AL565390.1	GI:12916718	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi,		
AUTHORS	Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo:		
TITLE	Li, W. B., Gruber, C., Jesse, J. and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 Evry cedex - France		

D	b	291	CGGGGAGATGCTGGGGCGGCTGGCTGTGTGGACCGCCGCCGGCCGAAGCTGAGCCTCTTTCG	350
OY		246	TgTgcTcogTgccttcTgtgcggcttTgcqTcaTcaacccgagccaggacgTgtga	305
D	b	351	TGTGTCTCCCTGCCTTCCTGTGGCCGGCTTTGCCGTATCAACCGGGCCAGAGGTGTGA	410
OY		306	TgctctcTggggggcgccgcctcTcaaccgacttgagcctgaggtTgttgtctccatgtgcc	365
D	b	411	TCTCTGTGGGGGGCGCCTTCCTCACCGGCGCTGGCCTTGCGGTGGCTCCCTAAGTGGCC	470
OY		366	cggTclaaTclccgaaatcgcTtaaccagacagTcccggggtTgtcTgcctcTgtgc	425
D	b	471	CGGTCACTACTCCGGAATCCGCTACCAGCAAGTCCGGGGGTTGCTCGGCTCTGTGTGC	530
OY		426	agcTaatagtcTgcTgcTgcgcattcTcct--ggcctaacttgagagcttggtgtgagtgaTg	483
D	b	531	AGCTAATAGTGTGCTGCGTGGCATCTCTCTTGGCCTTACTGTGCGAGGCTGGGTGTGAATG	590
OY		484	ggcgTgcgTgcgtgtgtgtgtgtgtgcTgcg-TgcTccccTccctcatgtcTgttcatgtct	542
D	b	591	GCGCTGGCGGTGTGTGTCTGGCTGGCTTCCCCCTTCCTCAATGCTGTTCTATGTGTCT	650
OY		543	tcaTcccccagaaccccgagcttcTcTgtcTgTaaTcaagacaaggcgccagagggccat--cg	599
D	b	651	TGATCGCCAGAACCCCGAGACTCTCTCTTACTACAGACAGGGCCAGAGGCATATGGCCG	710
OY		600	ccctcTgcgttcTgtgtgggTccTccagcagagTcTgtggaagaaccccccatcTgggtcTgagc	659
D	b	711	CTGGGGTTCCTTTGTGTGAGCTCCGACAGCGGTGGGAACAACCCCCTCATGCGGGGTGAGC	770
OY		660	aagagcttacct--ggcctcTgtgtggagccggacatctcaaaccccttatcatc	715
D	b	771	AGAGCTTTCACCTTGAGCCTTGTGTGGGAAGCCCGCATCTACAACCCCTTCATCATCT	828
RESULT	7			
LOCUS	BC771736	879 bp	mRNA	EST
DEFINITION	602720404r1 NIH_MGC_97 Homo sapiens CDNA clone IMAGE:4837297 5', mRNA sequence.			
ACCESSION	BC771736			
VERSION	BC771736.1	GI:14082389		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
TITLE	1 (bases 1 to 879)			
JOURNAL	NIH-MGC http://mhc.nci.nih.gov/			
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)			
	Contact: Robert Strausberg, Ph.D. Email: cga@ps-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihrai Toshiyuki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10769 row: h column: 02 High quality sequence stop: 765.			
FEATURES				
SOURCE	Location/Qualifiers			
	1..879			

[illegible]

	RESULT	11	
D8	Locus	BG281777	
	Definition	602403292ZP1 NIH_MGC_20 Homo sapiens CDNA clone IMAGE:4544686 5'	EST 21-FEB-2001
	Accession	BG281777	
	Version	BG281777.1	GI:13030703
	Keywords	EST.	
	Source	Homo sapiens	
	Organism	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Reference	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	Authors	1 (bases 1 to 838)	
	Title	NIH-MGC http://mgc.ncl.nih.gov/.	
	Journal	National Institutes of Health, Mammalian Gene Collection (MGC)	
	Comment	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabds@remail.nih.gov Tissue Procurement: ATCC/DCTD/DRP CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLML) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLM at: image.lnl.gov Plate: LNCMI228 row: g column: 23 High quality sequence stop: 784.	
	Features	Location/Qualifiers 1..838 <code>/organism="Homo sapiens"</code> <code>/db_xref="taxon:9606"</code> <code>/clone="IMAGE:4544686"</code> <code>/clone_lib="NH.MGC.20"</code> <code>/tissue_type="melanotic melanoma"</code> <code>/lab_host="DH10B (phage-resistant)"</code> <code>/note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcORI/XhoI sites using the following 5' adaptor: GGGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)." </code>	
	Base Count	90 a 306 c 282 g 158 t 2 others	
	Origin		
	Query Match	33.0%; Score 599; DB 11; Length 838:	
	Best Local Similarity	96.3%; Pred. No. 6,4e+83;	
	Matches	699; Conservative 0; Mismatches 16; Indels 11; Gaps 8;	
OY	6	gcgagcgccggcttcctcgc-gccttcgcgcgtgcccttgaggccaactacgttcgcg	64
D8	111	GCGGCCGCGCCTTCCTTGCCGTCCCTTCGCGCGCTGCCCCCACTCAGCTTCGCG	170
OY	65	tctgc-gctcgcctaagaattcccggaccacctgaagtcaaa-ggcgcggccccgg	122
D8	171	TTCGGTGCTCGACTCACATCCC CGGCCATTCCCTTAAGCTGAAGAACCAGGCCCGG	230
OY	123	ccecgcg-gccttgaaagcacgcgcgcacccctcgtgatcggtagctgtaacctggatgcc	181
D8	231	COCGCGTCGCTGAGACGACGCCCGCCCHCCTGCGTTCCGGGGCTGCTGACCTTCGCTGCC	290
OY	182	gcgcgcgggggaagtgcctgggcgcgtcgtcatgtgacgcgcgcggcgaccaactgaacctc	241
D8	291	GCGCGCGGGGAGTAGCTGCGCGCGCTGCGCTGGTGACCCGCGCGGCGCAACCTGACCCTC	350
OY	242	ttagcgtctccgtaccctcttgtcgcgcgtttgcgcgttatcataaacgcgcgccccagatcgt	301
D8	351	TTCGTGTGCTCCGTCCTCATCTGCGCGCGCTTGGCGGTATCATCCGCGGCCCATGAGACGTG	410

QY	302	tgatgtcgtcgtgaggggcccgtccctccaccagcctcgtgcgtgtgttgccctccatg	361
Db	411	TTGATGCTGCTGCGGGGGCCGCTCTCTACCCGGCTTGCGGTGTGGCTCTCCAGTG	470
QY	362	gcccgtctacatctaccgaatcgcctaccaccagcagtcgggggtgtgtcgtccctgt	421
Db	471	GCCCGGCTACATCTCGGAATGCGNCTACCCAGCAGCTCCGGGGTTGCTGGCTCTCTGT	530
QY	422	gtgcagctaagtgtcgtcgtcgtgcgcacccctcctgctactgtgcagcgtgtgtgtgc	481
Db	531	GTGCGAGCTAAAGTGTGCTGCTGGCATCTCTGTGCTACCTGTGCGAGGCTTGGTGTGAG	590
QY	482	tggcgtcgtcgtcgtcgt	541
Db	591	TGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT	650
QY	542	ttcatgcccgagaccgccgcgtctcgtctgtcgtacatcagacagcgcgcagagagccat	598
Db	651	TTTCATGCCCGAGAACCGCGCTTCTCTGTGATCAGCAGAGGCGCAGAGGCGCATGGCC	710
QY	599	gcccgtcgtcgtcgt	658
Db	711	GCCCTGCGGTCTCTGTGCGCTCCGAGAGGGCTGTGGAGAGA-CCCCCAATCGGGGCTGAG	769
QY	659	cagagcttcacccgtgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt	718
Db	770	CAGAACCTTTCACCTGTGGCTTAGTGTGG--AGCCCGGATCTTACAG-CTTTCATCATCGGA	826
QY	719	gtctcc 724	
Db	827	GTCTCC 832	
RESULT	12		
LOCUS	B114765	801 bp	EST
DEFINITION	602861314P1 NIH_MGC_17	Human sapiens cDNA clone	IMAGE:5020563 5',
ACCESSION	B114765		
VERSION	B114765.1	GI:14565666	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euthera; Primates; Carnivora; Homiidae; Homo.		
TITLE	1 (bases 1 to 801)		
JOURNAL	NIH-MGC http://mgi.ncl.nih.gov/		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgsds@mail.nih.gov		
	Tissue Procurement: ATCC		
	CDNA Library Preparation: Ling Hong/Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Inocyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LICM1837 row: 1 column: 04		
	High quality sequence stop: 783.		
FEATURES	Location/Qualifiers		
SOURCE	1..801		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5020563"		
	/clone_id="NIH_MGC_17"		
	/tissue_type="rhabdomyosarcoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;		
	Site_2: XhoI; cDNA made by oligo-dT priming.		
	directionally cloned into EcoRI/XhoI sites using the		
	following 5' adaptor: GGCAACGAG(G). Site-selected >500bp		

for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

Query Match	27.5%	Score 499.2	DB 11	Length 801
Best Local Similarity	87.9%	Pred No. 1.4e-67		
Matches 568	Conservative	0	Mismatches 73	Indels 5
				Gaps 2

[illegible]

RESULT	13
BE986058	
LOCUS	BE986058 664 bp mRNA EST 05-OCT-2000
DEFINITION	UI-M-CG0P-beu-e-08-0-UI.s1 NIH-BMAP Ret4_S52 Mus musculus cDNA clone
	UI-M-CG0P-beu-e-08-0-UI 3', mRNA sequence.

JOURNAL
MEDLINE
COMMENT
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Chin, H

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mestf@mail.nih.gov
Oligo-dt track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA library preparation: M.B.
Soares lab clone distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENERICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements
Seq primer: M13 Forward
POLY(A)-NO.

FEATURES	SOURCE	Location/Qualifiers
1..664		
/organism="Mus musculus"		
/strain="C57BL/6J"		
/db_xref="taxon:10090"		
/clone="U1-M-CG0p-beu-e-08-0-UI"		
/clone_lib="NIH_BMAP_Ret4_S2"		
/lab_host="DH10B (Life Technologies)"		
/note="Vector: pT73b-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The NIH_BMAP_Ret4_S2 library is a subcloned library, ultimately derived from mouse retina tissue libraries at various stages of development. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.utoro.ca .		
TAG_SEQ=None	Found"	
75 a	238 c	207 g 142 t 2 others

[illegible]

[illegible]

Db	219	TGCGCCTGGAGCTACAG6CTCCCGGCATCCTAGCCTGCAGCGCGCGGCCGCCGCC	278
Oy	126	cgcgccttgagcaacgcccgcgccctcctcgtgttcggggcctgctcgtgaccttgggtgcgcgg	185
Db	279	CGCGCCTGGAGCACAG-CGCGGCCTCCTGGTTGGGGCTTGTGAGACCTGGGTGCCGG	337
Oy	186	cggggggaagtcttgaggcgctgcgtggttgagccgccccgagctgaagcctctgc	245
Db	338	CGGGGGGAAGTCGTGGCGGCCTGCTGTGTGACC GCCCGGCCAAGCTGAGCCTTTGC	397
Oy	246	tgtgtcgcgtgcacctcgtgtagcgccgttcgcgcatcacccgccccagagcgtgtgaa	305
Db	398	TGTGCTCCGTCGCCCTTGCTGAGCGCGGCTTTGCGTCAACCGCGGCCAGAGCTGTGGA	457
Oy	306	tgtctctggggggccgcgccttcaccgcgcgttcggcgtggttgttctccctagttggcc	365
Db	458	TGCTCTCTGGGGGGCGCCCTTCACCGGGCTTGCGCGCTTGCTCCCTAGTG6CCC	517
Oy	366	cqgtclaca-ctccgaaaaatcgacctaacccagcaqtccgggggtgtctgctgcctgtg-t	423
Db	518	CGGTCTACTATTCTCCGAATTCGCCTAACCCAGCAGTCCGGGGGTGCTGGCTCTGTGTT	577
Oy	424	gcagctaattgltcgtcgtlcgcgalctcctcgtgacctactgtgcagcgttggtgtcgtga	483
Db	578	GCAGCTAAATGGTCGTCGTCGGCATCTACTAGCGCTACCTGAGCGAGCGTGAGTGTGAG	637
Oy	484	ggcgtgcgtgcgtgtgtcgtggtgcgtgccccctccctc	523
Db	638	GCGTGCTGCTGCTGTGCTGGCTGCTGCCCCCTCCTCCTC	677
RESULT	15		
LOCUS	AM955452	651 bp	mRNA
DEFINITION	EST367522	MAGE resequences,	MAGD Homo sapiens cDNA, mRNA sequence.
ACCESSION	AM955452		
VERSION	AM955452.1	GI:8145135	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 651)		
JOURNAL	Heide,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt		
COMMENT	, I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and Quackenbush,J. Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray Unpublished (2000) Contact: John Quackenbush The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel.: 301 838 3528 Fax: 301 838 0208 Email: johngnetigr.org Plate: 79 Seq primer: Reverse. Location/Qualifiers 1..651 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="MAGE_resequences,"MAGD" /note="Vector: pBluescriptSkm"		
BASE COUNT	146 a	185 c	169 g
ORIGIN	151 t		
Query Match	26.3%	Score 477:	DB 10: Length 651:
Best Local Similarity	97.6%:	Pred. No. 3.8e-64:	
Matches 494:	Conservative	0: Mismatches 11:	Indels 1: Gaps 1
Oy	1307	ttctctgtcccttgaaactaaagaagaaagactctggaacaataacacagcccatlgtgaagg	1366

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 15:01:01 : Search time 4552.73 Seconds
(Without alignments)
3757.652 Million cell updates/sec

Title: US-09-516-493-9
Perfect score: 1037
Sequence: 1 tgcgcgcgcctctagaacta.....aaaaactcgcggggggcgc 1037

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	756.6	73.0	2087	6 AX076669	AX076669 Sequence
2	756.6	73.0	2087	10 RN0245935	AJ245935 Rattus no
3	745.2	71.9	2189	10 AB033418	AB033418 Rattus no
4	483.6	46.6	2072	6 AX076671	AX076671
5	483.6	46.6	2072	10 MM0245936	AJ245936 Mus muscu
6	387.8	37.4	1843	10 AF232061	AF232061 Mus muscu
7	321.6	31.0	1490	10 MM017802	V17802 Mus muscu
8	274.4	26.5	1012	4 AF321324	AF321324 Bos tauru
9	272.6	26.3	1508	6 HSA17801	V17801 Homo sapien
10	271	26.1	2080	6 AX179740	AX179740 Sequence
11	268.8	25.9	1873	6 HSA245937	AJ245937 Homo sapi
12	268.8	25.9	2217	6 AX076667	AX076667 Sequence
13	125.4	12.1	225370	9 AL445222	AL445222 Human DNA
14	111.2	10.7	2011	6 AX076779	AX076779 Sequence
15	109.8	10.6	378	6 AX072418	AX072418 Sequence
16	109.8	10.6	1580	9 HSA17803	V17803 Homo sapien
17	109.8	10.6	2487	9 HSA011372	AJ011372 Homo sapi
18	108.2	10.4	1541	6 AX076677	AX076677 Sequence
19	99	9.5	42845	2 AC017247	AC017247 Drosophi
20	99	9.5	297385	3 AE003540	AE003540 Drosophi
21	92.6	8.9	1461	6 AX191507	AX191507 Sequence
22	88.8	8.6	185413	2 AC010025	AC010025 Drosophi
23	80.6	7.8	127914	2 AC093093	AC093093 Oryza sat
24	80.6	7.8	147706	8 AC083945	AC083945 Oryza sat
25	80	7.7	175506	3 AC008004	AC008004 Drosophi
26	77.2	7.4	2487	4 RABPCT	M21747 Rabbid gluc
27	76	7.3	177816	2 AC017643	AC017643 Drosophi
28	74.2	7.2	2533	4 BOVGGLUT1	M60448 Bos tauru
29	73.6	7.1	677	9 HSA296345	AJ296345 Homo sapi
30	73.2	7.1	2798	3 LEIDITRA	M85072 Leishmania
31	72.4	7.0	2230	4 SSGUTP	X17058 Sus scrofa
32	72.4	7.0	2473	10 MUSGLUTRN	M22998 Mouse facil
33	71.2	6.9	142114	8 AC073166	AC073166 Oryza sat
34	71	6.8	4521	9 AB056798	AB056798 Macaca fa
35	71	6.8	42415	9 AC001644	AC001644 Genomic s
36	71	6.8	44762	9 AC002325	AC002325 Homo sapi
37	69.2	6.7	2544	10 MUSST1A	M23384 Mouse glucu
38	69.2	6.7	198735	8 ATCHRTV44	AL161544 Arabidops
39	69.2	6.7	4126	6 HSA28H201	Z97341 Arabidops
40	68.6	6.6	4126	9 HS28H201	AL137188 Novel hum
41	68.6	6.6	4167	9 AF321240	AF321240 Homo sapi
42	68.6	6.6	4396	9 AF248053	AF248053 Homo sapi
43	67.6	6.5	2571	10 RATGLUTRN	M13979 Rat brain g
44	66.2	6.4	1578	8 ATH24968	AJ24968 Arabidops
45	66.2	6.4	1852	9 AF070544	AF070544 Homo sapi

ALIGNMENTS

RESULT 1
AX076669 LOCUS AX076669 2087 bp DNA
DEFINITION Sequence 3 from Patent WO0104145.
ACCESSION AX076669
VERSION AX076669.1 GI:12711200
KEYWORDS
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 2087)

1 (bases 1 to 2087)
AUTHORS Thorens,B., Ibberson,M. and Uldry,M.
TITLE Glutx polypeptide family and nucleic acids encoding same
JOURNAL Patent: WO 0104145-A 3 18-JAN-2001;
University of Lausanne (CH)
Location/Qualifiers

FEATURES

1..2087
/organism="Rattus sp." source

```

CDS
/db_xref="taxon:10118"
30..1466
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28496.1"
/db_xref="GI:12711201"
/translation="MSPEDPQETQPLLRSPGARPGRRVFLATFAALGSLPSFGL
GYSSPAIPSLRRTAPPALRLGDTAASWFGAVVTLGAAGVGLGMLDRGRKLSLL
CTVPEVTGFAVITARDVWMLLGGRLITGLACGASLAPVYISLAIYAVRGLLSGC
VOLAVVTGILLAVAGVLEWMLAVIGCPPTLMLLMCYMPETRPFLHOHOYOA
MAALRFMGSEGMERPPVGAEHOGFOLAMLRPVHKLPIITGLAMPOLSGVNAI
MEYANTIFDEAKFKDSSLASVYGIIOVLFYVAALIMDRGRKLLALSGLVIFSM
SAFGYFKLTQSGPNSSHVGLVPIISAEPADVHLGLAMLVGSMCLFTAGFVWGMP
IPMLMSEIFPLHKGAVTGCVLTNMFMAFLVTKFENSIMEILRPYGAFWLTAFCI
LSVLTFTFVEPTKGRTEQJTAHFEGR"
BASE COUNT      355 a      673 c      580 g      479 t
ORIGIN
Query Match      73.0%; Score 756.6; DB 6; Length 2087;
Best Local Similarity 90.8%; Pred. No. 4,3e-159;
Matches 899; Conservative 0; Mismatches 20; Indels 71; Gaps 6;

55 ctgttgccatctccgcagagcctgtcgtatgttaccctgggctgtgctgtgctgtgta 114
|||||
1098 ctgttgccatctccgcagagcctgtcgtatgttaccctgggctgtgctgtgctgtgta 1157
|||||
115 ggcagatgtcctctcaacgtgtgttgcagtaagctcggggagaccatccctgtgctc 174
|||||
1158 ggcagatgtcctctcaacgtgtgttgcagtaagctcggggagaccatccctgtgctc 1217
|||||
175 ctatgtcagagatctcctctgcacataaagggtgtgtaccaggtctgtgtctc 234
|||||
1218 ctatgtcagagatctcctctgcacataaagggtgtgtaccaggtctgtgtctc 1277
|||||
235 accaactgtctacgtcctctcgttgacaaaggatlttaagcgcatacagagatctc 294
|||||
1278 accaactgtctacgtcctctcgttgacaaaggatlttaagcgcatacagagatctc 1337
|||||
295 agaacctacggcctctctgtcctacgcgtcctctgtatctcctcagcgtctcttca 354
|||||
1338 agaacctacggcctctctgtcctacgcgtcctctgtatctcctcagcgtctcttca 1397
|||||
355 ctacacttctcctcagactaaaggcagactctggaacaaatcacagccatctga 414
|||||
1398 ctacacttctcctcagactaaaggcagactctggaacaaatcacagccatctga 1457
|||||
415 gacggtgagcagaccctctctgtcgtcagcgcctcgaagctgtgctgtgcttca 474
|||||
1458 gacggtgagcagaccctctctgtcgtcagcgcctcgaagctgtgctgtgcttca 1517
|||||
475 aaaggagtgagctcagtgacacacaglttgagcccaaggagccctcctactcctca 534
|||||
1518 aaaggagtgagctcagtgacacacaglttgagcccaaggagccctcctactcctca 1577
|||||
535 atttcggggcagcctctgttcacatactcaaccagattccacacatgagcttccaca 594
|||||
1578 atttcggggcagcctctgttcacatactcaaccagattccacacatgagcttccaca 1636
|||||
595 tcttgagactcmtgnagcctcgtcgcacacacagacatt----- 633
|||||
1637 tcttgagactcmtgnagcctcgtcgcacacacagacatt----- 1695
|||||
634 -----tgcggtcctcgtgct 648
|||||
1696 cttaaccacagcgtctccttcttaagactaaagcactgaagacagacactcgtcgctc 1755
|||||
649 ctatgtcctcgtcgtgacatcttggtgtgtgtgtcctaaagcactgcccatacctca 708
|||||
1756 ctatgtcctcgtcgtgacatcttggtgtgtgtgtcctaaagcactgcccatacctca 1815
|||||
709 ttgactgggggagtgagaaggagcttaggcacataagatttgggtccagaagaagtca 768
|||||

```

```

Db 1816 TTGACTGGGGGAGTGAAGAGGACTTAGCCACATTAACATTTGGGCTTCAGAAACAGGTCA 1875
Qy 769 ggtagctcagagagaagaagaatggttctgtctgtcgaacaaagctcttcagagt 828
|||||
Db 1876 GGTGAGTCCAGGAGAAAGAGANT-GTTCTTGTCTTGTCAACCACTCTTCTCAGAGT 1934
|||||
Qy 829 g-ccaaagcctcggatctacacttgggttgaagcagcttaccatcacttaccagttct 887
|||||
Db 1935 GCCCAGAGACCTCCGATTCACCTTGGGGTGTAGGACACTTAACCACTTACAGGTTCT 1994
|||||
Qy 888 ctccaactcagactgtgtcctcagtgctgcagatcatcagtcaccaggtctgttgatt 947
|||||
Db 1995 CTCCACTCTTCCCTGCTCAGTGTCTTGATCATTAAGTACACAGGCTTGTGAGATT 2054
|||||
Qy 948 cagaataaagaagcctcttcctcgtca 977
|||||
Db 2055 CAG-AAAATAAAGGCTCTTTCCTTCA 2083
|||||

RESULT 2
RNO245935 LOCUS
DEFINITION Rattus norvegicus mRNA for glucose transporter (GLUTX1 gene).
ACCESSION AJ245935.1 GI:7018604
VERSION 1
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2087)
AUTHORS Ibberson,M., Uldry,M. and Thorens,B.
TITLE GLUTX1, a novel mammalian glucose transporter expressed in the
JOURNAL J. Biol. Chem. 275 (7), 4607-4612 (2000)
MEDLINE 20138191
AUTHORS Ibberson,M.R.
TITLE 2 (bases 1 to 2087)
REFERENCE Direct Submission
AUTHORS Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
JOURNAL and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
CH-1005, SWITZERLAND
FEATURES
source location/Qualifiers
1..2087
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
30..1466
/gene="GLUTX1"
30..1466
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CAI75729.1"
/db_xref="GI:7018605"
/translation="MSPEDPQETQPLLRSPGARPGRRVFLATFAALGSLPSFGL
GYSSPAIPSLRRTAPPALRLGDTAASWFGAVVTLGAAGVGLGMLDRGRKLSLL
CTVPEVTGFAVITARDVWMLLGGRLITGLACGASLAPVYISLAIYAVRGLLSGC
VOLAVVTGILLAVAGVLEWMLAVIGCPPTLMLLMCYMPETRPFLHOHOYOA
MAALRFMGSEGMERPPVGAEHOGFOLAMLRPVHKLPIITGLAMPOLSGVNAI
MEYANTIFDEAKFKDSSLASVYGIIOVLFYVAALIMDRGRKLLALSGLVIFSM
SAFGYFKLTQSGPNSSHVGLVPIISAEPADVHLGLAMLVGSMCLFTAGFVWGMP
IPMLMSEIFPLHKGAVTGCVLTNMFMAFLVTKFENSIMEILRPYGAFWLTAFCI
LSVLTFTFVEPTKGRTEQJTAHFEGR"
BASE COUNT      355 a      673 c      580 g      479 t
ORIGIN
Query Match      73.0%; Score 756.6; DB 10; Length 2087;
Best Local Similarity 90.8%; Pred. No. 4,3e-159;
Matches 899; Conservative 0; Mismatches 20; Indels 71; Gaps 6;

55 ctgttgccatctccgcagagcctgtcgtatgttaccctgggctgtgctgtgctgtgta 114
|||||

```


Db 1910 CTGGCAGAGACTCTCCGGGGTCACCCCTGGGGCAGCCAGCTTAACCATTCACATTACAGGT 1969
Qy 885 tctctcaactctcagctggtctcagtgctcgtgatcatcagtcacaggtctgtgag 944
Db 1970 TCTTTCTGGCTCTTCTTGCGCTCCGCTGCTCGCTGATTAAGCAGCATATCTTTGTTGAG 2029
Qy 945 ttccagaaaaataaagcctcttcctcgttcaaaaaa 988
Db 2030 TTTTCAG-GAATTAAGCCTTACTGTTCAAAAAAACA 2072

RESULT 5
MMU245936 2072 bp mRNA ROD 18-FEB-2000
LOCUS Mus musculus mRNA for glucose transporter (GLUTX1 gene).
DEFINITION AJ245936
VERSION AJ245936.1 GI:7018606
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE house mouse.
ORGANISM Mus musculus.
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2072)
Ibberson, M., Uldry, M., and Thorens, B.
TITLE GLUTX1, a novel mammalian glucose transporter expressed in the
JOURNAL central nervous system and insulin-sensitive tissues
REFERENCE J. Biol. Chem. 275 (7), 4607-4612 (2000)
2 (bases 1 to 2072)
Ibberson, M.R.
AUTHORS Direct Submission
TITLE Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
JOURNAL and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
CH-1005, SWITZERLAND

FEATURES
Source location/Qualifiers
1..2072
/organism="Mus musculus"
/db_xref="taxon:10090"
21..1454
/gene="GLUTX1"
21..1454
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CAB75719.1"
/db_xref="GI:7018607"
/translation="MSPEDPQETQPLLREPEARTPRGRVPLASFALGPIPLFGAL
CTSPFAISLRRTAPPALRLGNAASWFGAVVILGAAGGILGGWLLDSGRKLSLL
CTVPVTGFAVITTAARDVWMLLGGRLTLGACGASLVAAPVISEIAYPAVRLLDSC
VOLAVVTGILLAVYAGWLEWMLAVLGCVPPTLMLLNCYMPETPRPLTQHOYOA
MALRFLWSEGEWEPYGAHOGFOLRLRPGIYKPLIIGISLWFOOLSGVNAI
MFYANSIFPEAKFKDSSLASVIGIIOVFTVAALIMRAGRRLTLASGVIMVSM
SAFGYFKLTOSLPSNSHVLGPIAAEPDVQVGLAMLVGSMCLFIAGFAAGMPI
PMLMSEIFPLHVKYATGICVLTNNFMAFLVTKFNSVMEMLRPYGATWLTAAFCAL
SVLPLTVVPEKGRLEQVTAHFEGR"

BASE COUNT 351 a 673 c 584 g 464 t
ORIGIN

Query Match 46.6%; Score 483.6; DB 10; Length 2072;
Best Local Similarity 75.3%; Pred. No. 5, 1e-98;
Matches 756; Conservative 0; Mismatches 161; Indels 87; Gaps 9;

Db 1206 CTATGTCAGAGACTTCCTTCCTGCAATGTCAGAGGTGTGGCTACCGGATCTGTGTCTC 1265
Qy 235 accaactggttcatcagcctcttcgttgagcaagaagatlttaacgatcatgagatctc 294
Db 1266 ACCAAGTGGTTCATGCGCTTTCTAGTAGCAAAAGAGTTCAACAGCGTATGGAGTGTG 1325
Qy 295 agaccctacgagcgcctcttcgttcacgcgtgctctctglatccctcagcgtcttcacg 354
Db 1326 AGACCTTAAGCGTGGCTTCTGCGCTCACCGCTCTTCTGCGCTCTGAGTGTCTATTAC 1385
Qy 355 ctcaacttgccttcagacacaaagcagacgtctggaacaaalcaacagcccatltgag 414
Db 1386 CTGACCGTGTCCCTGAGACATAAGCGAGGACTGTGAAACAAGTACAGCCCATTTTC 1445
Qy 415 gacggtgaacgagccctcttcgttcacgttcagccctcagcgtgagctggttcag 474
Db 1446 GGACGATGACAGACCCCTTTCTGTGCTGGAGCCCGGAGCTGAGCTGCTTTGGGTTTCA 1505
Qy 475 aaagagtgagtgagct-----cagtgacacagtltgagcccaaggggccccctga 526
Db 1506 GAAGAGAGTGAGTGCGCTGTAACTAAGCCACACCTCAATTTGAG-CTGGAGAGCCCTGA 1564
Qy 527 ctctcagattccgggccaagcttctcagatctcaacccagatltcaacacatgagct 586
Db 1565 CTGCTC--ACCTCAAGGCGCTCTTTGCCAGATCTCAACCCAGATTCCACCTGAGCC 1622
Qy 587 tcaccagatt--ctgagcgtctcgttcagcgtctcagacacagacatltcgagctctc 644
Db 1623 ACCAGACTCTAGTGTGCGGTCTCTGACGCTCTTCACACAGACATCTGAGGAGCTCT 1682
Qy 645 ggcctcagtgct----- 656
Db 1683 GCATCTCTACCTGACCCAGGTGTCTCTTTTAAGACTGAGGACAGCTGAAGGAGCTGAG 1742
Qy 657 -----ctgctgagcactcttgaggtgcttggtctcctaaagcagctccatccca 707
Db 1743 GCGTCTGCGCTCCGCTTCTCGGCTGGGCTCCGCTTAAGCAGCCGCTGATCTCA 1802
Qy 708 ctgactggggagtgagaaagagacttgacacataagatttggtctcagaaacaagctc 767
Db 1803 CTTGACTGCGGATCAGAAAGGACTTAGCC-----ATATGGGCTCAGAAACAGGTC 1855
Qy 768 aggtgagttccagaaagaagaatggtctctgtctgtcaacccaagctctcagag 827
Db 1856 AGGTGAGCGCCAGTAAAGAGAGAAATG-----TCTTGCCAACCAACCCCTCTCAGAG 1909
Qy 828 tgcacaaga-----cctcggatcaactcgtgggtgtagccagcttccatcaatcag 884
Db 1910 CTGGCAGAGACTCTCCGGGTCAACCCCTGGGCCAGCCAGCTTACCCATCTTAACAGGT 1969
Qy 885 tctctcaactcagctggtctcagtgctcgtgatactagtcacagcgtctgtgag 944
Db 1970 TCTTTCTGGCTCTTCTGCGGCTCCGCTGCTGATTAAGCAGCATATCTTTGTTGAG 2029
Qy 945 ttccagaaaaataaagcctcttcctcgttcaaaaaa 988
Db 2030 TTTTCAG-GAATTAAGCCTTACTGTTCAAAAAAACA 2072

RESULT 6
AF232061 1843 bp mRNA ROD 23-JUN-2000
LOCUS Mus musculus glucose transporter GLUT8 mRNA, complete cds.
DEFINITION AF232061
ACCESSION AF232061.1 GI:8671757
VERSION
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus.
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1843)
Carayannopoulos, M.O., Chi, M.M., Cui, Y., Plingssterhaus, J.M.,
McKnight, R.A., Mueckler, M., Devaskar, S.U., and Moley, K.H.

TITLE GLUT8 is a glucose transporter responsible for insulin-stimulated glucose uptake in the blastocyst

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (13), 7313-7318 (2000)

MEDLINE 20319023

REFERENCE 2 (bases 1 to 1843)

AUTHORS Moley, K.H., Carayannopoulos, M.O. and Cui, Y.

TITLE Direct Submission

JOURNAL Submitted (07-FEB-2000) OB/Gyn, Washington University, 4566 Scott Ave, St. Louis, MO 63110, USA

FEATURES

source 1.1843 Location/Qualifiers

5'UTR 49..1482 /organism="Mus musculus" /strain="129" /db_xref="taxon:10090" /product="glucose transporter GLUT8" /protein_id="AA78366.1" /db_xref="GI:8671738"

CDS 1..48 /codon_start=1

3'UTR 1483..1843

BASE COUNT 291 a 609 c 529 g 410 t 4 others

ORIGIN

Query Match 37.4%; Score 387.8; DB 10; Length 1843; Best Local Similarity 85.0%; Pred. No. 1.3e-76; Matches 472; Conservative 0; Mismatches 72; Indels 11; Gaps 3;

55 ctgtgtcccatctccgagagcctgtgactgattcaccctgggctgctgctgctgta 114
|||||
1114 CTGTGCGCCATCGCGGAGGAGCTGTGATGTCAGTGGGACTGGCTGCTGCTGTA 1173
|||||
115 ggcagcagtgctctctcactgctgctgctgctgctgctgctgctgctgctc 114
|||||
1174 GGCAGCATGTGCTCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 1233
|||||
175 ctcatgtcaagaatcttccctctgacatcaagggtgctgctgctgctgctgctc 224
|||||
1234 CTCTATGTCAGAGATCTTCCTCTGTCATGTCAAGGCTGCTGCTGCTGCTGCTC 1293
|||||
235 accaactggtlcatgagccttctctgctgctgctgctgctgctgctgctgctc 294
|||||
1294 ACCAAGTGTGCTTCTTCTAGTACCAAGAGTTCAGGCGCTCATGAGAGATGCTC 1353
|||||
295 agagccttagagagccttctgctcaccgctgctgctgctgctgctgctgctc 354
|||||
1354 AGAGCCTTAGAGCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 1413
|||||
355 ctcaactgttcccttagaactaaagcagagccttgaacaatacagcccaattgag 414
|||||
1414 CTGACCGTGTCTCCTCAGACTAAAGCAGAGACTCTGGAACAATCAGAGCCCTTC 1473
|||||
415 gacggtgacagccttctctgctgctgctgctgctgctgctgctgctgctgctc 474
|||||
1474 GACGATGTGACAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 1533
|||||
475 aaagagtgagtgagct-----cagtgacacagctttagaccaggggcccctga 526
|||||
1534 GAAGGATGTGAGTGCTTAACTAAGCCACACACTGAGCTTGAAGGCTGAGGCCCC 1593
|||||
527 ctctctagaattccgggacagcttctgacagatcaaaccaatacaacaatagct 586
|||||
1594 CTCCTC-ACCTCAGGGCCCTTCTTGCCTCAGATCTCAACCCAGATTC-CACTG 1650
|||||

QY 587 tcaccagatctcag 601
|||||
Db 1651 CCACCAGACTCTGAG 1665

RESULT 7

LOCUS MMU17802

DEFINITION Mus musculus mRNA for glucose transporter 8 (GLUT8 gene).

ACCESSION Y17802

VERSION Y17802.1 GI:7688219

KEYWORDS glucose transporter 8; GLUT8 gene.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1490) Doege, H., Schurmann, A., Bahrenberg, G., Brauers, A. and Joost, H.G. GLUT8, a novel member of the sugar transport facilitator family with glucose transport activity J. Biol. Chem. 275 (21), 16275-16280 (2000)

JOURNAL 2029367

MEDLINE 2 (bases 1 to 1490)

AUTHORS Joost, H.G.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-1998) H. Joost, Institute of Pharmacology and Toxicology, Technical University Aachen, Wendlingweg 2, D-52057 Aachen, FRG

FEATURES

source 1.1490 Location/Qualifiers

1.1490 /organism="Mus musculus" /db_xref="taxon:10090" /tissue="testis" /gene="GLUT8" /gene="GLUT8" 26..1459 /gene="GLUT8" 26..1459 /codon_start=1 /product="glucose transporter 8" /protein_id="CA889815.1" /db_xref="GI:7688220"

gene 26..1459

CDS 26..1459

BASE COUNT 222 a 499 c 438 g 331 t

ORIGIN

Query Match 31.0%; Score 321.6; DB 10; Length 1490; Best Local Similarity 87.8%; Pred. No. 8.5e-62; Matches 351; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

55 ctgtgtcccatctccgagagcctgtgactgattcaccctgggctgctgctgctgta 114
|||||
1091 CTGTGCGCCATCGCGGAGGAGCTGTGATGTCAGTGGGACTGGCTGCTGCTGTA 1150
|||||
115 ggcagcagtgctctctcactgctgctgctgctgctgctgctgctgctgctc 174
|||||
1151 GGCAGCATGTGCTCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 1210
|||||
175 ctcatgtcaagaatcttccctctgacatcaagggtgctgctgctgctgctgctc 234
|||||
1211 CTCTATGTCAGAGATCTTCCTCTGTCATGTCAAGGCTGCTGCTGCTGCTGCTC 1270
|||||
235 accaactggtlcatgagccttctctgctgctgctgctgctgctgctgctgctc 294
|||||

DB	1271	ACCAACTGGTTCATGGCCTTCTACTGACACCAAGAAGTTACGACGGCTCATGAGATGTGCTC	1330
OY	295	agaccctlaagcgccctcttgctgtacacgcgtgctctctgtatctccagcgtctcttcaag	354
DB	1331	AGACCTTACGGGCTTCTTGCTGCTACCCCTGCTTCTGCGCTTCAGTCTCTATTTCACA	1390
OY	355	ctcacccttgctccctcgagactaaagcgagactctggagacaaatcacgcccatttggg	414
DB	1391	CTGACCGTTGCTCCCTGAGAGCTAAAGGACAGACTCTGGAACACAGTACACGCCATTTCGAG	1450
OY	415	gacggtgacgagacccctctctgtgactgtgcaaccctgaagc	454
DB	1451	GGACGATGACAGACCCCTTCTCTGCTGCGAGACCCCGGACG	1490
RESULT	8		
AF321324			
LOCUS	AF321324	1012 bp	mRNA
DEFINITION	Bos taurus glucose transporter 8 (Glut8) mRNA, partial cds.		
ACCESSION	AF321324		
VERSION	AF321324.1	GI:14582715	
KEYWORDS			
SOURCE	cow.		
ORGANISM	Bos taurus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
	Bovidae; Bovinae; Bos.		
	1 (bases 1 to 1012)		
REFERENCE	Augustin, R., Navarrete-Santos, A. and Fischer, B.		
AUTHORS	Direct Submission		
TITLE	Submitted (13-NOV-2000) Anatomy and Cell Biology, Faculty of		
JOURNAL	Medicine, Martin Luther University Halle, Grosse Steinstrasse 52,		
	Halle 061097, Germany		
FEATURES	Location/Qualifiers		
source	1..1012		
	/organism="Bos taurus"		
	/db_xref="taxon:9913"		
gene	<1..1012		
	/gene="Glut8"		
CDS	<1..1005		
	/gene="Glut8"		
	/codon_start=1		
	/product="glucose transporter 8"		
	/protein_id="AAK69606.1"		
	/db_xref="GI:14582716"		
	/translation="1SETIAYPEVGRILGSCVOLMYVYTGILATYLAGWYLEWRMLAVLDC		
	CVASFMILMCFMPEMPRRLISQHKHOEAMAMOLFIMGVYOGNEEPPLGQHOEPIFYH		
	AQLRRPGYKPIITIGISLMAFQOLSGNANVFYETIIFEEKFRDSSLASVYGVTOV		
	LEFATNALIMDRGRRLLLTSGVVMWFSTAFQTYRFLTEGGSNSHDVLPALVSA		
	EAADTNGLAWLAGNMCLFTAGFAVAGSPFLPMLMSIEFLPHKGVATGVCVLTNNW		
	MAFLVTKEFSSLMLEVLRPYGAFLMFAFCIFGVFLTACVPETKGTLEQITLAFHEGR		
BASE COUNT	153 a	323 c	302 g
ORIGIN		234 t	
Query Match	26.5%:	Score 274.4:	DB 4: Length 1012:
Best Local Similarity	83.6%:	Pred. No. 2.9e-51:	
Matches 311:	Conservative	0: Mismatches	61: Indels
		0: Gaps	0:
OY	48	gacagagctgtgtgccatctccgacagagcgtgctgattcaccctgagctgtgactgagct	107
DB	630	GGACCTCCCGGCGCTGCTCATGATGAGAGCCCGCATACCAACGATGGGCTGGCTGGCT	689
OY	108	gagctgtaggcagcattgtgctctcttcatcttgcgtgtgttgcagttagcctggagccatccc	167
DB	690	GGCGGTGGGCAACATGTGCTTCATCGCCCGCTTCGCTGGCTGGGCGCCCATCC	749
OY	168	ctggctcccatgctacaagatcttccctctgacatacaaggtgtggtctaacggagctcty	227
DB	750	GTGGCTCTCATATGTAGAGATCTTCCCTCTGCATGTCAAGGGCGTGGCCACCGGGCTGTG	809
OY	228	tgtctcaccacacttgctcattgaccttctctgtgtgaccaaaggttlaacagcatcatgta	287

[illegible]

Accession	Sequence	Position
OY 117	caactgtgccccttcacacgcgtgtgtttgcagtagcctgaggagaccatccccttgacct	176
Db 1154	CAGCATGTGCGCTTTCATACGCCGGCTTGCGGCTGGGCTGGGGGCCCATCCCTGGCTCT	1213
OY 177	calgtcagaagatctccctctgcacatcaagaagtgtgtgtctaacgycgtcgtgltccctac	236
Db 1214	CATGTCAAGAGATCTTCCCTCTGCATGTCACAGGGCGTGGCGACAGCATCTCTGCTTCAC	1273
OY 237	caactgtgtctatgacctcttcctgtgcagccaagaagtttaacagcatcataggaatccctag	296
Db 1274	CACATGGCTCATGGCCTTTCTTCGTGACCAAGSAGATTACAGCCTCATGGAGGCTCTTAG	1333
OY 237	aacctaacggcgacctctgtgctctcaacgcgtgcctctctgtatccctcaagcgtcctcttca	356
Db 1334	GCCCATATGAGACCTTCTGGCTTGGCTCCGCCCTTCTTGACATCTTCAATGCTTCACATTT	1393
OY 357	caactcttctccctagactaaagacagagactctggaacaaatcaacagccatttgagg	415
Db 1334	TTCTGTGTCCCTGAACATTAAAGGAAGAACTCTGGAAATAATTCACAGCCCATTTTGAG	1452

RESULT	10			
AX179740				
LOCUS	AX179740	2080 bp	DNA	PAT
DEFINITION	Sequence 28 from Patent WO0146258.			06-AUG-2001
ACCESSION	AX179740			
VERSION	AX179740.1	GI:15132104		
KEYWORDS				
SOURCE				
ORGANISM	human.			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2080)			
	Baughn, M.R., Burford, N., Au-Young, J., Lu, D.A., Yang, J., Reddy, R., Lal, P., Hillman, J.L., Azimzai, Y., Yue, H., Nguyen, D.B., Yao, M.G., Gandhi, A.R., Tang, Y.T. and Khan, F.A.			
TITLE	Transporters and ion channels			
JOURNAL	Patent: WO 0146258-A 28 JUN-2001; Incyte Genomics, Inc. (US)			
FEATURES	Location/Qualifiers			
Source	1..2080			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/note="Incyte ID No: 1416107CB1"			
BASE COUNT	300 a 688 c 632 g 460 t			
ORIGIN				

Query Match	26.1%	Score 271;	DB 6;	Length 2080;
Best Local Similarity	84.7%	Pred. No. 1.8e-50;		
Matches 304; Conservative	0;	Mismatches 55;	Indels 0;	Gaps 0

[illegible]

Db 1413 GTTCTGTGTCCCTGAACCTAAAGGAAGACTCTGTGACCAACATCAGAGCCCATTTTGAGG 1471

RESULT	11
HSA245937	
LOCUS	
DEFINITION	HSA245937 1873 bp mRNA
DEFINITION	Homo sapiens mRNA for glucose transporter (GLUTX1 gene).
ACCESSION	AJ245937
VERSION	AJ245937.1 GI:7018305
KEYWORDS	glucose transporter; GLUTX1 gene.
SOURCE	human.

REFERENCE AUTHORS TITLE	JOURNAL MEDLINE REFERENCE AUTHORS TITLE
1 (bases 1 to 1873) Iberson, M., Uldry, M. and Thorens, B. GLUTX1, a novel mammalian glucose transporter expressed in the central nervous system and insulin-sensitive tissues J. Biol. Chem. 275 (7), 4607-4612 (2000)	2 (bases 1 to 1873) Iberson, M.R. Direct Submission Submitted (02-SEP-1999), Iberson M.R., Institute of Pharmacology and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne, CH-1005, SWITZERLAND

FEATURES	source	location/Qualifiers
gene		1..1873 /organism="Homo sapiens" /db_xref="taxon:9606"
CDS		4..1437 /gene="GLUTX1" 4..1437 /gene="GLUTX1" /codon_start=1 /product="glucose transporter" /protein_id="CA875702.1" /db_xref="GI:7018306"
BASE COUNT		279 a 621 c 573 g 400 t
ORIGIN		SVLTETTCVCEVLEIKGKTLTQITTAPEGR

Query Match Similarity	25.9%	Score	268.8	DB	9	Length	1873
Best Local Similarity	82.5%	Pred.	No. 5.7e-50				
Matches 320; Conservative	0	Mismatches	67	Indels	1	Gaps	1

[illegible]

```

Db 1311 GCCCTATGACGCTTCTGCGCTTCCGCTTTCGACATCTTCAGTGTCTTTTCACCTTT 1370
Qy 357 cacccttgcctcctagactaaagcaggaactcttgaacaaatcacaagcca-tttgaagg 415
Db 1371 GTTCTGTCTCCCTGAATAATTAAGGAAGACCTCTGGAACAAATCAAGCCCATTTTGAAGG 1430
Qy 416 acggtgaagagacccctcttctgtactgg 443
Db 1431 GCGATGACAGCCACTCATTTAGGAGATGG 1458

RESULT 12
AX076667 2217 bp DNA PAT 06-FEB-2001
LOCUS AX076667
DEFINITION Sequence 1 from Patent WO0104145.
ACCESSION AX076667
VERSION AX076667.1 GI:12711198
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2217)
AUTHORS Thorens,B., Ibberson,M. and Uldry,M.
TITLE Glutx polypeptide family and nucleic acids encoding same
JOURNAL Patent: WO 0104145-A 1 18-JAN-2001;
University of Lausanne (CH)
FEATURES
Source 1..2217
Location/Qualifiers
1..2217
/organism="Homo sapiens"
/db_xref="taxon:9606"
348..1781
/note="unlabeled protein product"
/codon_start=1
/protein_id="CAC8495.1"
/db_xref="GI:12711198"
/translation="MTPEDEPTNPPLGPPGSAAPRGRRVLAFAAAGLPLSFGAL
GYSPPAIPSLQRAAPAPRLDDAAASWFGAVNLGAAGAGVLGMLVDRAGRKLSLL
CSVPVAFVAVITAAQDVMMLGLRLTGLAGVASLVAPEYI5E1APVARGLLGSC
VOLMVVGLILAYLAGVLEMLAVLGVCPSPMLLMCFMPEPTPRELLTQHRORA
MALRLFMSEGDQEDPRICAGROSFALLRQPIKPIIGVSIAMFOOLSGVNAAM
FYAEITFEKPKFKDSSLASVAVGVIOVLTVAALINDRAGRRLILGSCVAVWFST
AFGAVKRLTGGGPNSSHVAISAPVSAQPDVASGLAMLVGNMCLIFAGAVAGWPI
PMLMSEIFPLHVKVATGICVLTNMLMAFLVTKFSSLMELVRLPYGAFVLASAFICIF
SVLTFLFCVPEIKGTLEOITAEHGR"
BASE COUNT 320 a 672 g 475 t
ORIGIN
Query Match 25.9%; Score 268.8; DB 6; Length 2217;
Best Local Similarity 82.5%; Pred. No. 5.8e-50;
Matches 320; Conservative 0; Mismatches 67; Indels 1; Gaps 1;

```

```

Qy 357 cacccttgcctcctagactaaagcaggaactcttgaacaaatcacaagcca-tttgaagg 415
Db 1715 GTTCTGTCTCCCTGAATAATTAAGGAAGACCTCTGGAACAAATCAAGCCCATTTTGAAGG 1774
Qy 416 acggtgaagagacccctcttctgtactgg 443
Db 1775 GCGATGACAGCCACTCATTTAGGAGATGG 1802

RESULT 13
AL445222 225370 bp DNA PRI 24-APR-2001
LOCUS AL445222
DEFINITION Human DNA sequence from clone RP11-356B19 on chromosome 9, complete
ACCESSION AL445222
VERSION AL445222.9 GI:13810082
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 225370)
AUTHORS Laird,G.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Apr 26, 2001 this sequence version replaced gi:13277497.
COMMENT During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em., EMBL; Sw.,
SWISSPROT; Tr., TREMBL; Mp., MORPEP; Information on the MORPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-356B19 is from the library RP11-1.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
left end of clone RP11-373J8 is at 96439 in this sequence. The true
right end of clone RP11-373J8 is at 9980 in this sequence.
FEATURES
Source 1..225370
Location/Qualifiers
1..225370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-356B19"
/clone_lib="RP11-1.2"
23455..23821
/note="Single clone region. Assembly confirmed by
restriction digest data."
65019..65283
/note="Single clone region. Sequence from reads from a
short insert library derived from a single pUC clone.
Restriction digest data confirm the assembly."
225123..225370
misc_feature
misc_feature

```

/note="Single clone region. Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."

BASE COUNT 56329 a 57092 c 57267 g 54682 t

ORIGIN

Query Match 12.1%; Score 125.4; DB 9; Length 225370;
Best Local Similarity 64.8%; Pred. No. 1.3e-17;
Matches 186; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 23 ggtatccccgggctgctgaatcgcgaagcgtgtgtccatccgcgaagcctgtgt 82
Db 61097 GGATCCGGGACCCACAGCCACAGCGGAGCGTGGCTCTGCTCCACAGCAAG 61156
QY 83 atgtacacgtgggctgctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 142
Db 61157 CTGAGAGAGCCAGCCCTGCGAGCCCTGATGCTCTTGGCTGTCTGCTCCACAGGCT 61216
QY 143 ttgcagtaagcttggagaccatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 202
Db 61217 TTGCGGTGGGCTGGGGCCCATCCCTGGCTCTCATGTACAGATCTTCCCTGATG 61276
QY 203 tcaaggt 262
Db 61277 TCAAGGGCTGTGCGAGCAGCATCTGCTGTCTCACCACTGCGTCATGGCTTCTGTGA 61336
QY 263 ccaaaagtttaacacatcatgtatcctgaagaccctgaagcgcgc 309
Db 61337 CCAAGAGATTACAGACCTCATGTGTGAGGCGAGGCTCCACAGCACC 61383

RESULT 14
AX076679 2011 bp DNA PAT 06-FEB-2001
LOCUS Sequence 13 from Patent WO0104145.
DEFINITION AX076679
ACCESSION AX076679.1 GI:12711210
VERSION
KEYWORDS
SOURCE Rattus sp.
ORGANISM Rattus sp.
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2011)
AUTHORS Thorens,B., Ibberson,M. and Uldry,M.
TITLE Glutx polypeptide family and nucleic acids encoding same
JOURNAL Patent: WO 0104145-A 13 18-JAN-2001;
University of Lausanne (CH)
FEATURES
Source location/Qualifiers
1..2011
/organism="Rattus sp."
/db_xref="taxon:10118"
7..1518
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28501.1"
/db_xref="GI:12711211"
/translation="MOEPLRAEGIDYDTPPEAPSPERKAGALONRRVFLATFPA
VLGNPFGYALYVMSVPIPAKRRSDPALRDKIOASFGSVFTLGAAGSLAMLLN
DLGRKLSINSAVPSAIGTALMAGARLMLLGRMTGTAGGTLTAACIPVYSEIA
PVGVRGAGATPQMLAAVFGSLTVALGILLPMRLAAVAGEPVAVMILLSPMNSPR
FLSKSRDEALQALIMLRADSEVMEFEQJODNVRROSSVNAEAPVPRPILI
TVLMRFLQOLGTIPILVYLOTIFDSTSVLPQODAAIVAVRLSVLIAAVMDLA
GRKVLVVASIMFVANLTGLVYOLVPRITLPSSTVETVITLGTETQPPRAAVYRL
ILATLMLFMGVMGMPITWMLSEVPLPRAGVASGLCVLVSMLAEVLTFTPL
AVNAFGIADVPFFPSAICLLSLFTGCCVPTKRSLEQIEAFHTRMSRP"

BASE COUNT 337 a 621 c 593 g 459 t 1 others
ORIGIN

Query Match 10.7%; Score 111.2; DB 6; Length 2011;
Best Local Similarity 62.0%; Pred. No. 1e-14;

Matches 176; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 127 ctcttcatgcgtgttttcagtaagcttggagaccatccctggtctctcatgtcagag 186
Db 1204 CTCTCATATTATGGGCTATGTCATGGCTGGGGCCCATACCTGCTCCATGTTGTAG 1263
QY 187 atctccctctgcacatacaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 246
Db 1264 GTTCTCCCCCTGGTGGTCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1323
QY 247 atggccttctgtgtaccaaagatttaacagatcatgtgagatctcgaacctcagagc 306
Db 1324 ACAGCCTTCGCTCTCACTAAGTACTTCTGTCGACGATGAATGCCCTTCGAGGCT 1383
QY 307 gctcttgcctacagcgtctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 366
Db 1384 CTTTCTTCTTCTTCTTCTGCGCATCTGCTGTGAGCTCTTCTTCTTCTTCTTCTTCT 1443
QY 367 cctgaagactaaagcagagacttgaacaatacaagccattt 410
Db 1444 CCTGAGACCAGGCGCGCTGCTGAGAGCATGTGAGGCTTCTT 1487

RESULT 15
AX072418 378 bp DNA PAT 25-JAN-2001
LOCUS Sequence 2890 from Patent WO0102568.
DEFINITION AX072418
ACCESSION AX072418
VERSION AX072418.1 GI:12582769
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 378)
AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.,
Kassam,A., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D.,
Lamsom,G., Dmanac,R., Gikenjakov,R., Dmanac,S., Dickson,M.,
Labat,I., Leshkowitz,D., Kita,D., Garcia,Y. and Strache-Grain,B.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0102568-A 2890 11-JAN-2001;
CHIRON CORPORATION (US) ; HYSO, INC. (US)
FEATURES
Source location/Qualifiers
1..378
/organism="Homo sapiens"
/db_xref="taxon:9606"

Query Match 10.6%; Score 109.8; DB 6; Length 378;
Best Local Similarity 62.6%; Pred. No. 1.7e-14;
Matches 171; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 127 ctcttcatgcgtgttttcagtaagcttggagaccatccctggtctctcatgtcagag 186
Db 31 CTCTCATATTATGGGCTATGTCATGGCTGGGGCCCATACCTGCTCCATGTTGTAG 90
QY 187 atctccctctgcacatacaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 246
Db 91 GTTCGCTCCCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 150
QY 247 atggccttctgtgtaccaaagatttaacagatcatgtgagatctcgaacctcagagc 306
Db 151 ACCGCTTCGCTCTACCAAGTCTTCTGCGACGTGTGAGACACTTGGGCTCCAGGTG 210
QY 307 gctcttgcctacagcgtctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 366
Db 211 CTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 270
QY 367 cctgaagactaaagcagagacttgaacaatacaagc 399
Db 271 CCGGAGACCAGGAGGCGCTGCTGAGAGCATGTGAGGCTTCTT 303

Search completed: February 13, 2002, 15:02:23
Job time: 8070 sec

QY	55	ctggctcccatctccgcagagagcctctgtcatgttcaaccctggggcttgccctggctggctgtga	114
Dp	1086	ctggctcccatctccgcagagagcctctgtcatgttcaaccctggggcttgccctggctggctgtga	1145
QY	115	ggcagcatgtgcctcttcatcagctctgtgtttcttgcagttgagctgtggagccatccctggagct	174
Dp	1146	ggcagcatgtgcctcttcatcagctctgtgtttcttgcagttgagctgtggagccatccctggagct	1205
QY	175	ctcatctcagagatctctccctctgcacatccaagaggtgtgtgtctaacccgggtctctgtctcc	234
Dp	1206	ctcatctcagagatctctccctctgcacatccaagaggtgtgtgtgtctaacccgggtctctgtctcc	1265
QY	235	accaactggttcatgtgcctctctctgtgtgaccacaaagatttaaacagatcatatgtgagatctcc	294
Dp	1266	accaactggttcatgtgcctctctctgtgtgaccacaaagatttaaacagatcatatgtgagatctcc	1325
QY	295	agaccctcagcgccctctctgtctccacccgtctgcctctgtatctccagcgctctcttcaacg	354
Dp	1326	agaccctcagcgctctctctgtctccacccgtctgcctctgtatctccagcgctctcttcaacg	1385
QY	355	ctcacctctgtccctctgagacttaagacagagactctgtgaacaacatcacagcccatcttgag	414
Dp	1386	ctgacgcctgtccctctgagacttaagacagagactctgtgaacaacatcacagcccatcttgag	1445
QY	415	gagcgctgagcagaccctctctgtgactctgcagccctcgagctttagctgtgctctgggtttca	474
Dp	1446	gagcgctgagcagaccctctctgtgactctgcagccctcgagctttagctgtgctctgggtttca	1505
QY	475	aaaaggatctgagatgtgcct-----cagctgacacagatttgagccagggggcccccctga	526
Dp	1506	aaaaggatctgagatgtgcct-----cagctgagacagatttgagccagggggcccccctga	1564
QY	527	ctctccagattctccgggcagcctctgttccagatctcaacccaagattccaacacatctagct	586
Dp	1565	ctctcc--acctcagggccctctcttccacagatctcaacccaagattccaacacatctagct	1622
QY	587	tcacacagatt--ctgagagcctctgtagcagctctgcagacacacagacacatttggcgctct	644
Dp	1623	acccagactctgagatgtgcgtctctgcagcctgtcttccacacagagacacatctgagagcgtct	1682
QY	645	ggctctagtgct-----	656
Dp	1683	ggctctctccacatctgacccagatgtctctctcttctaagactgagcctgagagacatctgag	1742
QY	657	-----ctggctgggcatctcttgggtgtcttggcttgaagcactgcgcatactctca	707
Dp	1743	ggctctctggctcccggttctctctggctgggtgtcccggtctcttaagcagcgctctgtatctca	1802
QY	708	cttcaactggggagatgaagaaagagactcttaagcacaataagatttgggtctcagaagaaagctc	767
Dp	1803	cttgaactggggagatctcagaagaaagagactcttaagcacaataagatttgggtctcagaagaaagctc	1855
QY	768	aggtgtgactcagagagaagaaagagatgtctctgtctgtctcaacaagatcctctctcagag	827
Dp	1856	aggtgtgactcagagagaagaaagagatgt-----ctcttccaacaagagcctctctcagag	1909
QY	828	tgccaaagaa---ctccggagatctaaccttggggtttgacagcgttaccatctactaacagt	884
Dp	1910	ctggcgacagagactctccggggtacacctctgggcgcagccagcttaccatctactaacagt	1969
QY	885	ctctctcaactctcaagctctgtctccagctgtctccctgtgagatctcaactgaagcagctctgttga	944
Dp	1970	ctctctcgctcttctctctgtgcctcgtgtccgtgtccatctgaagcacaacatactgtgtgag	2029
QY	945	tttcaagaanaaagagcctctcttcggttcaaaaaaanaana 988	
Dp	2030	tttcaag--gaaataaaagcctcttactgttcaaaaaaanaana 2072	

RESULT	3
AAD09552	
ID	AAD09552 standard; cDNA; 2080 BP.

XX	AA009552;	
AC		
XX		
DT	10-SEP-2001	(first entry)
XX		
DE	Human transporter and ion channel-1 (TRICH-1) CDNA.	
XX		
KW	Human; transporter and ion channel-1; TRICH-1; vaccine; cystic fibrosis;	
KW	gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy;	
KW	hypertension; angina; neurological disorder; asthma; bipolar disorder;	
KW	dementia; depression; Alzheimer's disease; epilepsy; mood; arrhythmia;	
KW	Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke;	
KW	Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy;	
KW	demyelinating disease; mental disorder; Schizophrenia; polyomyositis;	
KW	muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease;	
KW	dermatomyositis; diabetes mellitus; immunological disorder; psoriasis;	
KW	rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus;	
KW	sickle cell anaemia; Wilson's disease; infertility; Cushing's disease;	
KW	scleroderma; pulmonary artery stenosis; neutropenic Addison's disease;	
KW	malabsorption syndrome; hypercholesterolaemia; cancer; ss.	
OS	Homo sapiens.	
XX		
XX		
FH	Key	Location/Qualifiers
FT	CDS	46..1479
FT		/*tag= a
FT		/product= "Human TRICH-1 protein"
XX		
PN	WO200146258-A2.	
XX		
PD	28-JUN-2001.	
XX		
PF	22-DEC-2000; 2000WO-0355095.	
XX		
PR	23-DEC-1999; 99US-0172000.	
PR	14-JAN-2000; 2000US-0176083.	
PR	21-JAN-2000; 2000US-0177332.	
PR	28-JAN-2000; 2000US-0178572.	
PR	02-FEB-2000; 2000US-0179758.	
PR	10-FEB-2000; 2000US-0181625.	
XX		
PA	(INCY-) INCYTE GENOMICS INC.	
XX		
PI	Baughn MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy R, Lal P;	
PI	Hillman JL, Azimzal Y, Yue H, Nguyen DB, Yao MG, Gandhi AR.	
PI	Tang YF, Khan FA;	
XX		
DR	WPI: 2001-418042/44.	
DR	P-PSDB: AAE04888.	
XX		
PT	Novel human transporter and ion channel proteins useful for treating	
PT	and preventing transport, neurological, muscle and immunological	
PT	disorders	
XX		
PS	Claim 5; Page 140-141; 160pp; English.	
XX		
CC	The present sequence is transporter and ion channel-1 (TRICH-1) CDNA.	
CC	TRICH is used as vaccine. TRICH is useful for treating a disease or	
CC	condition associated with decreased expression of functional TRICH,	
CC	such as transport disorder including amyotrophic lateral sclerosis,	
CC	cystic fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth	
CC	disease, Duchenne muscular dystrophy, angina and hypertension,	
CC	neurological disorders including Alzheimer's disease, amnesia, bipolar	
CC	disorder, dementia, depression, epilepsy, ischaemic cerebrovascular	
CC	disease, stroke, cerebral neoplasms, Pick's disease, Huntington's	
CC	disease and Parkinson's disease, demyelinating diseases, mental disorders	
CC	including mood, anxiety, Schizophrenia and seasonal affective disorder,	
CC	muscle disorder including cardiomyopathy, myocarditis, polyomyositis,	
CC	dermatomyositis, arrhythmias and asthma and immunological disorders	
CC	including AIDS, adult respiratory distress syndrome (ARDS), allergies,	
CC	anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's	
CC	syndrome, systemic lupus erythematosus and other diseases including	
CC	sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary	

CC artery stenosis, Grave's disease, Cushing's disease, Addison's disease,
 CC glucosyl-galactose malabsorption syndrome, hypercholesterolemia, cancers
 CC psoriasis and viral, bacterial, fungal, helminthic and protozoal
 CC infections. TRICH DNA is useful in gene therapy and in diagnostic
 CC purposes.

XX
 SQ Sequence 2080 BP; 300 A; 688 C; 632 G; 460 T; 0 other;

Query Match 26.1%; Score 271; DB 22; Length 2080;
 Best Local Similarity 84.7%; Pred. No. 8.2e-49;
 Matches 304; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 57 ggtgccatctcgcagagcctgtgattcaccctgggctgctgctgctgag 116
 DB 1113 ggcgcctgtctgcagacgctgtgatgccagctgggctgctgctgctgag 1172
 QY 117 cagcatgtcctctcctcctgctgttctgagtagctgggagccatccctgctcct 176
 DB 1173 cagcatgtcctcctcctcctcctgctgtgctgctgctgctcct 1232
 QY 177 catgcaagatctcctcctgcacatcaagggtgtgctacccgctgtctctcctac 236
 DB 1233 catgcaagatctcctcctgcacatcaagggtgtgctacccgctgtctcctac 1292
 QY 237 caactgttcattgctcttctgtgtgaccaaagatttaacagatcattgagatcctcag 296
 DB 1293 caactgttcattgctcttctgtgtgaccaaagatttaacagatcattgagatcctcag 1352
 QY 297 accctcagcgcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 356
 DB 1353 gccctatgagcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1412
 QY 357 caacttgcctcctgagactaaaggcagagctctgagacaatcagccatttgaggg 415
 DB 1413 gtctgtgtccctgaaactaaaggagagctctgagacaatcagccatttgaggg 1471

RESULT 4

AAF55865 standard; cDNA; 2217 BP.

XX
 AC AAF55865;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE Human GLUTX1 coding sequence.
 XX
 KW Human; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hypoglycaemia; ss;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200104145-A2.
 XX
 PD 18-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-IB01042.
 XX
 PR 14-JUL-1999; 99US-0143907.
 PR 27-AUG-1999; 99US-0151140.
 PR 23-FEB-2000; 2000US-0184285.
 PR 13-JUL-2000; 2000US-0616132.
 XX
 PA (UYLA-) UNIV LAUSANNE.
 XX
 PI Thorens B, Ibberson M, Uldry M;
 XX
 DR WP1: 2001-112615/12.
 DR P-PSDB; AAB66932.
 XX
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in

PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -

XX Claim 3; Page 70-71; 124pp; English.

XX The present invention relates to GLUTX proteins (AAF55865-AAF55871) and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is the coding sequence
 CC for human GLUTX1.

XX
 SQ Sequence 2217 BP; 320 A; 750 C; 672 G; 475 T; 0 other;

Query Match 25.9%; Score 268.8; DB 22; Length 2217;
 Best Local Similarity 82.5%; Pred. No. 2.5e-48;
 Matches 320; Conservative 0; Mismatches 67; Indels 1; Gaps 1;

QY 57 ggtgccatctcgcagagcctgtgattcaccctgggctgctgctgctgag 116
 DB 1415 ggcgcctgtctgcagacgctgtgatgccagctgggctgctgctgctgctgag 1474
 QY 117 cagcatgtcctcctcctcctcctgctgttctgagtagctgggagccatccctgctcct 176
 DB 1475 cagcatgtcctcctcctcctcctcctgctgttctgagtagctgggagccatccctgctcct 1534
 QY 177 catgcaagatctcctcctgcacatcaagggtgtgctacccgctgtctgtctcctac 236
 DB 1535 catgcaagatctcctcctgcacatcaagggtgtgctacccgctgtctgtctcctac 1594
 QY 237 caactgttcattgctcttctgtgtgaccaaagatttaacagatcattgagatcctcag 296
 DB 1595 caactgttcattgctcttctgtgtgaccaaagatttaacagatcattgagatcctcag 1654
 QY 297 accctcagcgcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 356
 DB 1655 gccctatgagcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1714
 QY 357 caacttgcctcctgagactaaaggcagagctctgagacaatcagccca-tttgaggg 415
 DB 1715 gtctgtgtccctgaaactaaaggagagctctgagacaatcagccca-tttgaggg 1774
 QY 416 acggtgagagcctccttctgtgactgg 443
 DB 1775 gcgagagagcctcctccttctgtgactgg 1802

RESULT 5

AAA44914/C
 ID AAA44914 standard; cDNA; 579 BP.

AAA44914;

DT 21-AUG-2000 (first entry)

DE Human secreted expressed sequence tag SEQ ID NO:1489.

XX Human; mouse; chicken; rat; secreted expressed sequence tag; SEQN;
 KW expressed sequence tag; EST; probe; chemokine; analgesic; hemostatic;
 KW immunomodulatory; haematopoietic; chemokine; analgesic; hemostatic;
 KW thrombolytic; antiinflammatory; cytoskeletal; antibacterial; antifungal;
 KW antiviral; antidiabetic; antiasthmatic; antiparkinsonian;
 KW anticancer; osteoprotective; neuroprotective; neurotropic; antiparkinsonian;
 KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;
 KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
 KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
 KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
 KW central nervous system disorder; Alzheimer's disease; stroke;
 KW Parkinson's disease; Huntington's disease; coagulation disorder;
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;


```

Db      1384 ccttctccttcctctgcgcatcgcttgctcaagccgtctctcacaggctgctgtag 1443
Oy      367 ccctgaacctaaagcagagactctggacaataacagaagccattt 410
       ||||| | | | | | | | | | | | | | | | | | | | | |
Db      1444 ccgagaccagggcgctgcctggagcagatcgagagccctctt 1487
RESULT 7
ID AAF67134
AC AAF67134 standard; cDNA; 378 BP.
XX
DT 09-APR-2001 (first entry)
DE Novel human polynucleotide, SEQ ID NO: 2890.
XX
KW Human: cytostatic; gene therapy; colon cancer; prostate cancer;
   breast cancer; lung cancer; cancer detection; ss.
XX Homo sapiens.
OS MO200102568-A2.
PN 11-JAN-2001.
PD 30-JUN-2000; 2000MO-USI8374.
PR 02-JUL-1999; 99US-0142310.
PX 02-JUL-1999; 99US-0142311.
PA (CHIR ) CHIRON CORP.
   (HYSE ) HYSEQ INC.
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Kravchenkov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Grain B;
DR WPI: 2001-091805/10.
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
PS Claim 9; Page 974; 1046bp; English.
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping OF
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX Sequence 378 BP; 45 A; 130 C; 117 G; 86 T; 0 other;
S0
Query Match          10.6%; Score 109.8; DB 22; Length 378;
Best Local Similarity 62.6%; Pred. No. 1.3e-14;
Matches 171; Conservative 0; Mismatches 102; Indels 0; Gaps
Oy      127 ctctcacagtgttgtagttgagcagtgtagtgagccatcccctggtctctcatgtcagag 186
       ||||||| | | | | | ||||| ||||| ||||| ||||| ||||| |||

```

Db	31	ccctacatcattggtctacgcgcgtgggctgtgggtgccatcaactgctctcatcattgtcagg	90
Qy	187	atcttcctctcgtcacatcaagggtgtgtgctaccggcgtctgtgtctccacacactgctc	246
Db	91	gtcttcgccttcgtcgccgttgccgttgccctcaggcctctgctgtgtgcacgcgtgc	150
Qy	247	atggcctctctcgtgaccacaagaagtttaacaagatcatgttgatctcagacctcagc	306
Db	151	acgcgccttcgtcccccaccaaagctctctccgtccagttgttagacactctggtccaggt	210
Qy	307	gcctctcgtctcaccacgcctctctctgtatccctcagctccctcttcitacagctacctt	366
Db	211	cccttcctctctcttcgcggccatctcgttggagcctctgtgttcaacggcctgtgtg	270
Qy	367	ccctgagactaaagcagagactctggaacaatc	399
Db	271	cccgagaccacaaggacggtctcctcgtgacagatc	303
RESULT	8		
AAA23461			
ID	AAA23461	standard; cDNA; 823 BP.	
XX	AAA23461;		
AC			
XX	19-JUN-2000	(first entry)	
DT			
DE		cDNA encoding human secreted protein vq1_1, SEQ ID NO:77.	
XX			
KW		Human; secreted protein; cancer; tumour; cardiovascular disorder;	
KW		blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;	
KW		infection; fungal; bacterial; viral; HIV; allergy; arthritis;	
XX		neurodegenerative disease; asthma; contraceptive; ss.	
OS		Homo sapiens.	
XX			
XX			
EH	Key	Location/Qualifiers	
FT	CDS	54..371	
FT		/*tag= a	
XX		/product= "Human secreted protein vq1_1"	
PN	WO200011015-A1.		
XX			
PD	02-MAR-2000.		
XX			
PF	24-AUG-1999;	99WO-US19351.	
XX			
PR	24-AUG-1998;	98US-0097638.	
PR	24-AUG-1998;	98US-0097659.	
PR	09-SEP-1998;	98US-0099618.	
PR	28-SEP-1998;	98US-0102092.	
PR	25-NOV-1998;	98US-0109978.	
PR	23-DEC-1998;	98US-0113645.	
PR	23-DEC-1998;	98US-0113646.	
XX	23-AUG-1999;	99US-0379246.	
XX			
PA	(ALPH-) ALPHAGEN INC.		
XX			
PI	Valenzuela D, Yuan O, Hoffman H, Hall J, Raple]ko P;		
XX			
DR	WPI: 2000-224657/19.		
DR	P-PSDB: AAY95019.		
XX			
PT	New secreted or transmembrane proteins and polynucleotides encoding		
PT	them, useful for treating neurodegenerative disorders, autoimmune		
PT	diseases and cancer -		
XX			
PS	Claim 86; Page 335; 357pp; English.		
XX			
CC	The invention relates to 40 human secreted proteins (AAY94981-Y95020),		
CC	and cDNA sequences encoding them (AAA23423-A23462). The secreted		
CC	proteins of the invention include those that are thought to be only		
CC	partially secreted, i.e., transmembrane proteins. The proteins of the		

CC invention may exhibit one or more activities selected from the following:
 CC cytokine activity; cell proliferation; differentiation; immune
 CC modulation; haematopoiesis regulation; tissue growth activity;
 CC activin/inhibin activity; chemotactic/chemokinetic activity; haemostatic
 CC and thrombolytic activity; anti-inflammatory activity; and tumour
 CC inhibition activity. The proteins may be administered to patients as
 CC vaccines, and the nucleotides may be used as part of a gene therapy
 CC regime. Diseases or conditions that may be treated using the proteins or
 CC nucleotides of the invention include autoimmune diseases: genetic
 CC disorders; haemophilia; cardiovascular diseases; cancer; bacterial,
 CC fungal and viral infections, especially HIV; multiple sclerosis;
 CC rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome;
 CC insulin dependent diabetes mellitus; and allergic reactions such as
 CC asthma and anaemia. They may also be used for treating wounds, burns,
 CC ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease and amyotrophic
 CC lateral sclerosis (ALS). Proteins with activin/inhibin activity may
 CC additionally be useful as contraceptives. Nucleic acid sequences of the
 CC invention may be used in chromosome mapping, and as a source of
 CC diagnostic primers and probes. The present sequence represents cDNA
 CC encoding one of the 40 proteins of the invention.

SO Sequence 823 BP: 170 A: 257 C: 233 G: 163 T: 0 other:

Query Match 10.6%; Score 109.8; DB 21; Length 823;
 Best Local Similarity 62.6%; Pred. No. 1.6e-14;
 Matches 171; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

OY 127 ccttcacatgcgtgttgcagtagcgtgagaccacccctgctcctcagtcagag 186
 DB 57 ccttcacatgcgtgttgcagtagcgtgagaccacccctgctcctcagtcagag 116
 OY 187 attccctcctcgcacatcaagggtgtgtgctaccgctgtgtgtctccacacgtgttc 246
 DB 117 gtcctgcctccctgcgtgcccgtggtgctgctcaggtcctgctgctgctgctgctc 176
 OY 247 atgctcctctcgtgagcaagaagtttaacagacatcattcctcagaccctcagcgc 306
 DB 177 accgctcctgctcctcaccagagctcctcgtcagcagtggtgagaccttgcgtccaggtg 236
 OY 307 gctctcgtgctaacgcgtgctcctgtatcctcagcgtccttcaagctcacccttgc 366
 DB 237 ccttctcctctcctcgcagcgtcctgctgtgagccttggttctcagcgtcgtgtg 296
 OY 367 cctgagactaaagcagagactctggaacaatc 399
 DB 297 ccgagaccaaaggagcgtccctgagcagatc 329

RESULT 9
 AAC80569
 ID AAC80569 standard; cDNA: 1577 BP.

AC AAC80569;
 DT 12-FEB-2001 (first entry)

DE Human secreted protein gene 39 SEQ ID NO:49.

XX Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;
 XX antiproliferative; cytostatic; cardiant; vasotrophic; cerebroprotective;
 XX neurotropic; neuroprotective; antibacterial; virucide; fungicide; cancer;
 XX optalmalological; autoimmune disease; hyperproliferative disorder;
 XX cardiovascular disorder; cerebrovascular disorder; wound healing;
 XX nervous system disorder; aging; chemotaxis; ss.

OS Homo sapiens.
 XX
 XX WO200058467-A1.
 XX
 XX PD 05-OCT-2000.
 XX

PF 22-MAR-2000; 2000MO-US07505.
 XX
 XX 26-MAR-1999; 99US-0126502.
 PR 17-DEC-1999; 99US-0172410.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM, Komatsoulis G;
 PI
 DR WPI: 2000-611712/58.
 DR P-PSDB: AAB45158.
 XX
 XX Nucleic acid molecules encoding human secreted proteins, used in
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 PT Parkinson's diseases and cancers -
 PS Claim 1: Page 371-372; 440pp; English.

CC polynucleotide sequences AAC80531-C80580 represent cDNA encoding human
 CC secreted proteins AAB45120-B45169. Sequences AAB45170-B45225 represent
 CC alternative polypeptides encoded by the genes, and amino acid sequences
 CC to which they are homologous. The genes and proteins have activities
 CC dependent on the tissues and cells in which they are expressed. Examples
 CC of their activities include immunosuppressive; antiarthritic;
 CC antirheumatic; antiproliferative; cytostatic; cardiant; vasotrophic;
 CC cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;
 CC fungicide; and optalmalological. The secreted proteins, polynucleotides,
 CC antagonists and agonists may be useful in treating, preventing and/or
 CC diagnosing diseases and disorders such as autoimmune disease
 CC e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms
 CC of the breast or liver, cardiovascular disorders e.g. cardiac arrest,
 CC cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous
 CC system disorders e.g. Alzheimer's disease, infections caused by
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
 CC The polypeptides can also be used to aid wound healing and epithelial
 CC cell proliferation, to prevent skin aging due to sunburn, to maintain
 CC organs before transplantation, for supporting cell culture of primary
 CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
 CC also be used as a food additive or preservative to increase or decrease
 CC storage capabilities. AAC80522-C80530 and AAB45119 represent sequences
 CC used in the isolation and characterisation of the genes and proteins of
 CC the invention.

SO Sequence 1577 BP: 352 A: 496 C: 445 G: 284 T: 0 other:

Query Match 10.6%; Score 109.8; DB 21; Length 1577;
 Best Local Similarity 62.6%; Pred. No. 1.9e-14;
 Matches 171; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

OY 127 ccttcacatgcgtgttgcagtagcgtgagaccacccctgctcctcagtcagag 186
 DB 291 ccttcacatgcgtgttgcagtagcgtgagaccacccctgctcctcagtcagag 350
 OY 187 attccctcctcgcacatcaagggtgtgtgctaccgctgtgtctcctcaccacgtgttc 246
 DB 351 gtcctgcctccctgcgtgcccgtggtgctcaggtcctgctgctgctgctgctc 410
 OY 247 atgctcctctcgtgagcaagaagtttaacagacatcagagatcctcagaccctcagcgc 306
 DB 411 accctcctgctcctcaccagagctcctcctcgtcagcagtggtgagcacttgcgtccagtg 470
 OY 307 gctctcgtgctaacgcgtgctcctgtatcctcagcagcgtccttcaagctcacccttgc 366
 DB 471 ccttctcctctcctcgtggtcgtcgtgtggtgagcctggtgttctcagcaggtcgtgtgtg 530
 OY 367 cctgagactaaagcagagactctggaacaatc 399
 DB 531 ccgagaccaaaggagcgtccctgagcagatc 563

RESULT 10
 AAF55870

PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 06-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 09-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 10-AUG-1999; 99US-0147935.
 PR 11-AUG-1999; 99US-0148171.
 PR 12-AUG-1999; 99US-0148319.
 PR 13-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 16-AUG-1999; 99US-0148684.
 PR 17-AUG-1999; 99US-0149368.
 PR 18-AUG-1999; 99US-0149175.
 PR 20-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151303.
 PR 01-SEP-1999; 99US-0151338.
 PR 07-SEP-1999; 99US-0151930.
 PR 10-SEP-1999; 99US-0152363.
 PR 13-SEP-1999; 99US-0153070.
 PR 15-SEP-1999; 99US-0153758.
 PR 16-SEP-1999; 99US-0154018.
 PR 20-SEP-1999; 99US-0154039.
 PR 22-SEP-1999; 99US-0154779.
 PR 23-SEP-1999; 99US-0155139.
 PR 24-SEP-1999; 99US-0155486.
 PR 28-SEP-1999; 99US-0155659.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 18-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159684.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160761.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160988.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match
 Best Local Similarity 7.4%; Score 76.6; DB 21; Length 1395;
 Matches 167; Conservative 0; Mismatches 129; Indels 3; Gaps 1;

QY 106 ctgctgtaagcagcctgaccccttcacgccttgctgacccaagagttacagcatg 165
 Db 1084 ctgctgctgtgtgtataatgtgtacatagatcgtcttcacagcagatgagcatg 1143
 QY 166 ccctgctccatagctacagagatctcccttcacatcaagagggtgctacacggcgc 225
 Db 1144 ccgtgggtgtatgtctgagatattcccaataaagaagtagcagagcatg 1203
 QY 226 tgtctccacacactgtgttcacgtcttctgtgacccaagagttacagcatg 285
 Db 1204 gcgacgctggtgactgtgttgagcgtggcgtgttcttaccattcacttcactc 1262
 QY 286 gagatccctcagacccctacgcgcctctgtgacacgcctgctctatccacgctc 345
 Db 1263 -gtcttgagctcttaccagacttctcctatlaacgtgcacacagcgtgacat 1320
 QY 346 ctltcacgcacactgttccctgagactaaagcagagctctggaacaatcacagc 404
 Db 1321 gtcttgcatctgcacatcgtcgtgagacaagaagagacatgtgagcagatccagc 1379

RESULT 13
 AAC39099 standard; DNA; 1683 BP.
 ID AAC39099;
 AC AAC39099;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23364.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128234.
 PR 16-APR-1999; 99US-0128714.
 PR 19-APR-1999; 99US-0129845.
 PR 21-APR-1999; 99US-0130077.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999;	99US-0132486.
PR 06-MAY-1999;	99US-0132487.
PR 07-MAY-1999;	99US-0132863.
PR 11-MAY-1999;	99US-0134256.
PR 14-MAY-1999;	99US-0134218.
PR 14-MAY-1999;	99US-0134219.
PR 14-MAY-1999;	99US-0134221.
PR 14-MAY-1999;	99US-0134370.
PR 18-MAY-1999;	99US-0134768.
PR 19-MAY-1999;	99US-0134941.
PR 20-MAY-1999;	99US-0135124.
PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135629.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 10-JUN-1999;	99US-0138847.
PR 14-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 16-JUN-1999;	99US-0139453.
PR 17-JUN-1999;	99US-0139492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 18-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140891.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 01-JUL-1999;	99US-0142154.
PR 06-JUL-1999;	99US-0142055.
PR 08-JUL-1999;	99US-0142350.
PR 09-JUL-1999;	99US-0142803.
PR 12-JUL-1999;	99US-0142920.
PR 13-JUL-1999;	99US-0142977.
PR 14-JUL-1999;	99US-0143542.
PR 15-JUL-1999;	99US-0143624.
PR 16-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144085.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144637.
PR 20-JUL-1999;	99US-0144884.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145085.
PR 22-JUL-1999;	99US-0145087.

PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145919.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 14-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.

Query Match	7.48;	Score 76.6;	DB 21;	Length 1683;
Best Local Similarity	55.9%;	Pred. No. 2.4e-07;		
Matches 167; Conservative	0;	Mismatches 129;	Indels 3;	Gaps 1;

[illegible]

RESULT	ID	Accession	Gene
14	AAC45298	AAC45298	Standard; DNA; 1685 bp.
XX	AC	AAC45298;	
XX	DT	18-OCT-2000	(first entry)
XX	DE	Arabidopsis thaliana	DNA fragment SEQ ID NO: 46007.
XX	XX	Hybridisation assay;	genetic mapping; gene expression control;
XX	KM	protein identification;	signal transduction pathway;
XX	XX	metabolic pathway;	promoter; termination sequence; ss.
OS	XX	Arabidopsis thaliana.	
PN	XX	EP1033405-A2.	
PD	XX	06-SEP-2000.	
FE	XX	25-FEB-2000;	2000EP-0301439.
XX	XX	25-FEB-1999;	99US-0121825.
XX	XX	09-MAR-1999;	99US-0123160.
XX	XX	23-MAR-1999;	99US-0123548.
XX	XX	25-MAR-1999;	99US-0125788.
XX	XX	29-MAR-1999;	99US-0126264.
XX	XX	01-APR-1999;	99US-0126785.
XX	XX	06-APR-1999;	99US-0127462.
XX	XX	08-APR-1999;	99US-0128234.
XX	XX	16-APR-1999;	99US-0128714.
XX	XX		99US-0129845.

PR	19-APR-1999;	9905-0130077
PR	21-APR-1999;	9905-0130449
PR	23-APR-1999;	9905-0130510
PR	23-APR-1999;	9905-0130891
PR	28-APR-1999;	9905-0131149
PR	30-APR-1999;	9905-0132048
PR	04-MAY-1999;	9905-0132407
PR	05-MAY-1999;	9905-0132484
PR	06-MAY-1999;	9905-0132485
PR	07-MAY-1999;	9905-0132487
PR	11-MAY-1999;	9905-0132863
PR	14-MAY-1999;	9905-0134256
PR	14-MAY-1999;	9905-0134218
PR	14-MAY-1999;	9905-0134221
PR	14-MAY-1999;	9905-0134370
PR	18-MAY-1999;	9905-0134768
PR	19-MAY-1999;	9905-0134941
PR	20-MAY-1999;	9905-0135124
PR	21-MAY-1999;	9905-0135312
PR	24-MAY-1999;	9905-0135353
PR	25-MAY-1999;	9905-0135629
PR	27-MAY-1999;	9905-0136021
PR	27-MAY-1999;	9905-0136392
PR	28-MAY-1999;	9905-0136782
PR	01-JUN-1999;	9905-0137222
PR	03-JUN-1999;	9905-0137528
PR	04-JUN-1999;	9905-0137502
PR	07-JUN-1999;	9905-0137724
PR	08-JUN-1999;	9905-0138094
PR	10-JUN-1999;	9905-0138540
PR	10-JUN-1999;	9905-0138847
PR	14-JUN-1999;	9905-0139119
PR	16-JUN-1999;	9905-0139452
PR	16-JUN-1999;	9905-0139453
PR	17-JUN-1999;	9905-0139452
PR	18-JUN-1999;	9905-0139454
PR	18-JUN-1999;	9905-0139455
PR	18-JUN-1999;	9905-0139456
PR	18-JUN-1999;	9905-0139457
PR	18-JUN-1999;	9905-0139458
PR	18-JUN-1999;	9905-0139459
PR	18-JUN-1999;	9905-0139460
PR	18-JUN-1999;	9905-0139461
PR	18-JUN-1999;	9905-0139462
PR	18-JUN-1999;	9905-0139463
PR	18-JUN-1999;	9905-0139750
PR	18-JUN-1999;	9905-0139753
PR	21-JUN-1999;	9905-0139817
PR	22-JUN-1999;	9905-0139899
PR	23-JUN-1999;	9905-0140353
PR	23-JUN-1999;	9905-0140354
PR	24-JUN-1999;	9905-0140635
PR	24-JUN-1999;	9905-0140823
PR	30-JUN-1999;	9905-0140991
PR	30-JUN-1999;	9905-0141287
PR	01-JUL-1999;	9905-0141842
PR	01-JUL-1999;	9905-0142154
PR	02-JUL-1999;	9905-0142055
PR	06-JUL-1999;	9905-0142390
PR	08-JUL-1999;	9905-0142803
PR	09-JUL-1999;	9905-0142820
PR	12-JUL-1999;	9905-0142977
PR	13-JUL-1999;	9905-0143342
PR	14-JUL-1999;	9905-0143624
PR	15-JUL-1999;	9905-0144005
PR	16-JUL-1999;	9905-0144085
PR	16-JUL-1999;	9905-0144085
PR	19-JUL-1999;	9905-0144425
PR	19-JUL-1999;	9905-0144331
PR	19-JUL-1999;	9905-0144332
PR	19-JUL-1999;	9905-0144333
PR	19-JUL-1999;	9905-0144334

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125789.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128274.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 99US-0143542.

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 22-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149920.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150366.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 26-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0157117.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 13:34:27 ; Search time 92.83 Seconds
(without alignments)
2529.976 Million cell updates/sec

Title: US-09-516-493-9

Perfect score: 1037

Sequence: 1 tggcgagccgtctagacta.....aaaacatcgagggggcc 1037

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 11328999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62.4	6.0	2503	1 US-08-472-934-3	Sequence 3, Appli
2	62.4	6.0	2503	1 US-08-472-934-11	Sequence 11, Appli
3	62.4	6.0	2503	1 US-08-323-460A-3	Sequence 3, Appli
4	62.4	6.0	2503	2 US-08-461-146C-3	Sequence 3, Appli
5	62.4	6.0	2503	2 US-08-461-146C-11	Sequence 11, Appli
6	62.4	6.0	2503	3 US-08-461-145C-3	Sequence 3, Appli
7	62.4	6.0	2503	3 US-08-461-145C-11	Sequence 11, Appli
8	61.8	6.0	1223	3 US-09-154-874-4	Sequence 4, Appli
9	60.8	5.9	1641	1 US-08-300-903A-8	Sequence 8, Appli
10	60.8	5.9	2085	2 US-08-668-128B-7	Sequence 7, Appli
11	60.8	5.9	2085	2 US-08-905-445-7	Sequence 7, Appli
12	60	5.8	2567	3 US-08-993-260-4	Sequence 4, Appli
13	59.8	5.8	1813	4 US-09-071-224-3	Sequence 3, Appli
14	59.2	5.7	1461	3 US-08-722-126A-4	Sequence 4, Appli
15	59.2	5.7	1461	5 PCT-US95-04258-4	Sequence 4, Appli
16	59.2	5.7	2409	4 US-09-293-322C-8	Sequence 8, Appli
17	58.6	5.6	500	4 US-08-818-112-101	Sequence 101, App
18	58.2	5.6	470	4 US-09-020-956-102	Sequence 102, App
19	58.2	5.6	470	4 US-09-030-607-102	Sequence 102, App
20	58.2	5.6	509	4 US-09-030-607-202	Sequence 202, App
21	58.2	5.6	5503	2 US-08-726-012B-1	Sequence 1, Appli
22	58	5.6	2605	2 US-08-680-395-4	Sequence 4, Appli
23	58	5.6	2880	4 US-09-115-954-3	Sequence 3, Appli
24	58	5.6	3842	4 US-09-115-954-7	Sequence 7, Appli
25	58	5.6	3912	4 US-09-115-954-1	Sequence 1, Appli
26	57.8	5.6	3089	1 US-08-472-934-5	Sequence 5, Appli
27	57.8	5.6	3089	2 US-08-323-460A-5	Sequence 5, Appli

28	57.8	5.6	3089	2 US-08-461-146C-5	Sequence 5, Appli
29	57.8	5.6	3089	3 US-08-461-145C-5	Sequence 5, Appli
30	57.8	5.6	5173	1 US-08-242-677-1	Sequence 1, Appli
31	57.6	5.6	688	6 5498694-3	Patent No. 5498694
32	57.2	5.5	1066	1 US-08-157-101A-4	Sequence 4, Appli
33	57.2	5.5	1074	4 US-09-357-251-23	Sequence 23, Appli
34	57.2	5.5	1315	4 US-09-721-822A-10	Sequence 10, Appli
35	57.2	5.5	1558	4 US-08-455-550-7	Sequence 7, Appli
36	57.2	5.5	2407	4 US-09-020-956-110	Sequence 110, App
37	57.2	5.5	3410	4 US-09-030-607-110	Sequence 110, App
38	57.2	5.5	3410	4 US-08-467-607-2	Sequence 2, Appli
39	57	5.5	1558	1 US-08-469-362-2	Sequence 2, Appli
40	57	5.5	1558	2 US-08-850-392-2	Sequence 2, Appli
41	57	5.5	1577	4 US-08-821-994-59	Sequence 59, Appli
42	57	5.5	1577	4 US-08-821-994-59	Sequence 59, Appli
43	56.8	5.5	1738	2 US-08-379-482A-2	Sequence 2, Appli
44	56.6	5.5	1601	1 US-07-968-971A-10	Sequence 10, Appli
45	56.6	5.5	1601	1 US-08-424-406-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-472-934-3
Sequence 3, Application US/08472934
Patent No. 5753446
GENERAL INFORMATION:
APPLICANT: JOHNSON, GARY L.
TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESS: Lahive and Cockfield
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,934
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/440,421
FILING DATE: 15-May-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/354,516
FILING DATE: 21-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 5,405,941
FILING DATE: 15-Apr-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,460
FILING DATE: 14-Oct-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/11690
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04178
FILING DATE: 15-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr., Esq.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: CPI-004DVC02
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2503 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 466..2325
US-08-472-934-3

Query Match
Best Local Similarity 66.2%; Score 62.4; DB 1; Length 2503;
Pred. No. 9.8e-06;
Matches 90; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1 tggagcgccgtctagactagtagatccccgggctgcaggaattcgacagagctggtg 60
|||||
DB 3 TGGGCGCCGCTTAGAAGTGTGATCCCCGGCTGCAGGAATTGCGACGAGGACGA 62

QY 61 cccatctcgcagagcctgtgatgtcacctggggctggtccctggtgctgtagcagc 120
|||||
DB 63 TCCAGCGCGCAGAGTCCGCCCTTCCGCTTCTCTCCGGTGGCGACGGGCGCCGG 122

QY 121 atgtgcctcttcacg 136
|||||
DB 123 GCCTTCCTTTTCATCG 138

RESULT 2
US-08-472-934-11
Sequence 11, Application US/08472934
Patent No. 5753446
GENERAL INFORMATION:
APPLICANT: JOHNSON, GARY L.
TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
NUMBER OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive and Cockfield
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,934
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/440,421
FILING DATE: 15-May-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/354,516
FILING DATE: 21-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 5,405,941
FILING DATE: 15-Apr-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,460
FILING DATE: 14-Oct-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/11690
FILING DATE: 14-Oct-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04178
FILING DATE: 15-Apr-1994

ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. Deconati, Jr. Esq.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: CPI-004DVCP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2503 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 466..2325
US-08-472-934-11

Query Match
Best Local Similarity 66.2%; Score 62.4; DB 1; Length 2503;
Pred. No. 9.8e-06;
Matches 90; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1 tggagcgccgtctagactagtagatccccgggctgcaggaattcgacagagctggtg 60
|||||
DB 3 TGGGCGCCGCTTAGAAGTGTGATCCCCGGCTGCAGGAATTGCGACGAGGACGA 62

QY 61 cccatctcgcagagcctgtgatgtcacctggggctggtccctggtgctgtagcagc 120
|||||
DB 63 TCCAGCGCGCAGAGTCCGCCCTTCCGCTTCTCTCCGGTGGCGACGGGCGCCGG 122

QY 121 atgtgcctcttcacg 136
|||||
DB 123 GCCTTCCTTTTCATCG 138

RESULT 3
US-08-323-460A-3
Sequence 3, Application US/08323460A
Patent No. 5854043
GENERAL INFORMATION:
APPLICANT: JOHNSON, GARY L.
TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
NUMBER OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: SHERIDAN ROSS & MCINTOSH
STREET: 1700 LINCOLN STREET, SUITE 3500
CITY: DENVER
STATE: CO
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/323,460A
FILING DATE: 14-Oct-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/049,254
FILING DATE: 14-Apr-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOVARIK, JOSEPH E.
REGISTRATION NUMBER: 33,005
REFERENCE/DOCKET NUMBER: 2879-1-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/863-9700
TELEFAX: 303/863-0223
INFORMATION FOR SEQ ID NO: 3:

Db 76 CTCA 79


```

? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/722,126A
? FILING DATE: 08-OCT-1996
? CLASSIFICATION: 536
?
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/US95/04258
? FILING DATE: 06-APR-1995
?
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: IL 109257
? FILING DATE: 08-APR-1994
?
? ATTORNEY/AGENT INFORMATION:
? NAME: BROWDY, Roger L.
?
? REGISTRATION NUMBER: 25,618
? REFERENCE/DOCKET NUMBER: PECTHT-1A
?
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 628-5197
? TELEFAX: (202) 737-3528
?
? INFORMATION FOR SEQ ID NO: 4:
?
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1461 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
?
? MOLECULE TYPE: CDNA
?
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 54..617
?
? US-08-722-126A-4

```

```

Query Match          5.7%; Score 59.2; DB 3; Length 1461;
Best Local Similarity 88.9%; Pred. No. 4.6e-05;
Matches 64; Conservative 0; Mismatches 8; Indels 0; Gaps 0.

QY 952 aaataaagagccttcctccgtcacaataaaaaaaaaaaaaaaaaaaaaaa 1011
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1369 AAAATAAAAAATACCTTTCTGCTAAAAAaaaaaaaaaaaaaaaaaaaaa 1428
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1012 aaaaaaaaaa 1023
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1429 AAAAAAAAAAAAA 1440

RESULT 15
PCT-US95-04258-4
; Sequence 4, Application PC/TUS9504258
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A DNA MOLECULE ENCODING A MAST CELL
; TITLE OF INVENTION: FUNCTION-ASSOCIATED ANTIGEN (MARA)
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDT AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04258
; FILING DATE: 06-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 109257

```

```

1 FILING DATE: 08-APR-1994
2 ATTORNEY/AGENT INFORMATION:
3 NAME: BROWDY, Roger L.
4 REGISTRATION NUMBER: 25,618
5 REFERENCE/DOCKET NUMBER: PCHT=1 PCT
6 TELECOMMUNICATION INFORMATION:
7 TELEPHONE: 202-628-5197
8 TELEFAX: 202-737-3528
9
10 TELEX: 248633
11
12 INFORMATION FOR SEQ ID NO: 4:
13 SEQUENCE CHARACTERISTICS:
14 LENGTH: 1461 base pairs
15 TYPE: nucleic acid
16 STRANDEDNESS: single
17 TOPOLOGY: linear
18 MOLECULE TYPE: CDNA
19
20 FEATURE:
21 NAME/KEY: CDS
22 LOCATION: 5..617
23
24 PCT-US95-04258-4

```

	Query Match	Best Local Similarity	Score	5.7%;	DB 5;	Length	1461;
	Matches	64; Conservative	0;	Mismatches	8;	Indels	0;
Qy	952	aaaaaaaaagcctcttcttcgtaaaaaaaaaaaaaaaaaaaaaaaaaa	1011				
Db	1369	AAATTTAAAAATATCCTTTTCTCAAAAAAAAAAAAAAAAAAAAAA	1428				
Qy	1012	aaaaaaaaaaaaa	1023				
Db	1429	AAAAAAAAAAAAA	1440				

```
Search completed: February 13, 2002, 13:34:39
Job time: 2806 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 13:37:33 ; Search time 2842.4 Seconds
(Without alignments)
3920.410 Million cell updates/sec

Title: US-09-516-493-9

Perfect score: 1037

Sequence: 1 tgcgcgcgcctctagaacta.....aaaaaccgcagggggggcc 1037

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 segs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: qb_est1:*
11: qb_est2:*
12: qb_hic:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	364.2	35.1	563	11	B1275901 UT-R-CX0-
C 2	364	35.1	538	11	BG672321 DRNCPD10
C 3	316.6	30.5	450	11	BF420038 UT-R-B12-
C 4	316.4	30.5	419	11	H34451
C 5	315.6	30.4	481	10	AA997295
C 6	312.2	30.1	923	11	BF120840
C 7	271	26.1	949	11	BG749509
C 8	268.2	25.9	546	10	BE231636
C 9	248.2	23.9	295	11	BF420257
C 10	231.6	22.3	426	11	BF551290
C 11	231.6	22.3	934	11	BE910478
C 12	231.4	22.3	712	11	BG290178

C 13	229	22.1	324	11	H34372
C 14	218.6	21.1	752	10	AL565390
C 15	199.6	19.2	796	13	AZ985752
C 16	193.4	18.6	631	11	BG066327
C 17	184.8	17.8	603	11	BG079541
C 18	182.2	17.6	235	11	BF146289
C 19	180.2	17.4	461	11	BF459354
C 20	179.2	17.3	496	10	AM248655
C 21	178.6	17.2	554	11	BG099902
C 22	178	17.2	420	11	BI202551
C 23	176.6	17.0	342	11	BI220862
C 24	176.2	17.0	440	11	BF774206
C 25	163.6	15.8	447	10	AA213079
C 26	162.6	15.7	409	10	AM783633
C 27	162.6	15.7	523	11	BG065950
C 28	162.4	15.7	917	11	BF140667
C 29	159.4	15.4	583	11	C77528
C 30	159	15.3	401	10	AA254369
C 31	157	15.1	327	11	BF742266
C 32	152.6	14.7	294	10	AV003143
C 33	150.4	14.5	595	11	BG079217
C 34	147.2	14.2	338	10	BB53574
C 35	145.4	14.0	314	10	BB272673
C 36	143.4	13.8	301	10	BB388688
C 37	143.4	13.8	318	10	BB164525
C 38	140.4	13.5	290	10	BB351479
C 39	139.8	13.5	529	10	AA790557
C 40	139.4	13.4	287	10	AV015118
C 41	138.8	13.4	324	10	BB503238
C 42	138.6	13.4	314	10	BB206085
C 43	138.2	13.3	257	11	C78846
C 44	137.8	13.3	275	10	AA616872
C 45	136.8	13.2	332	10	BB558923

ALIGNMENTS

RESULT 1
LOCUS B1275901/c
DEFINITION UT-R-CX0-bxh-h-03-0-UI-s1 UT-R-CX0 Rattus norvegicus cDNA clone
ACCESSION B1275901
VERSION B1275901
KEYWORDS B1275901.1 GI:14920290
SOURCE EST.
ORGANISM Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 563)

AUTHORS Bonaldo,M.F., Jenson,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized rat placenta pool library cDNA library preparation. M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com)
Seq primer: M13 Forward

Y	453	gcggagctgctcggagcttcaaaaagatlgagatggccctcagctggaccacagctttaggc	512
Db	301	gctgagctgcttccgggttccaaagagatgcagctgacgtgaccttactgacacacagcttgacc	360
Y	513	caggagccccctgaacccctcagatctccggcgccagctctgtccagatctcaccagatc	572
Db	361	caggggccccctgactctcagatctccgagctccggcagccttggctacgactcgaacca	419
Y	573	ccacacatgagcttccaccagatcttgagctctcgaagctctgcacacacagacatc	632
Db	420	ccacacacatgagcttccaccagatcttgagctctcgaagctctgcacacacagacatc	478
Y	633	ttgcggagctctc	645
Db	479	ttgcggagctctc	491
RESULT	3		
LOCUS	Bf420038	450 bp	mRNA
DEFINITION	U1-R-BJ2-bpm-c-04-0-01.s1 U1-R-BJ2 Rattus norvegicus cDNA clone		
ACCESSION	Bf420038		
VERSION	Bf420038.1	GI:11408027	
KEYWORDS	EST.		
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus		
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.		
TITLE	Normalization and subtraction: two approaches to facilitate gene		
JOURNAL	Genome Res. 6 (9), 791-806 (1996)		
MEDLINE	97044477		
COMMENT	Contact: Soares, MB Program for Rat Gene Discovery and Mapping University of Iowa 451 Eckstein Medical Research Building Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: msoares@blue.weeg.uiowa.edu The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized heart library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) The following repetitive elements were found in this cDNA sequence: 1-21, >At_richlow_complexity Seq primer: M13 Forward POLYA=yes.		
FEATURES			
SOURCE	Location/Qualifiers		
	1..450		
	/organism="Rattus norvegicus"		
	/strain="Sprague-Dawley"		
	/db_xref="taxon:10116"		
	/clone="U1-R-BJ2-bpm-c-04-0-01"		
	/clone_lib="U1-R-BJ2"		
	/lab_host="DH10B (Life Technologies)"		
	/note="Vector: pRT73D-pac (Pharmacia) with a modified polylinker: Site_1: Not I; Site_2: Eco RI; The U1-R-BJ2 library is a subtracted library derived from the following tissues: heart, atrium at 15 dpc, ventricle at 16.5 dpc, atrium at 16.5 dpc, ventricle at 13 dpc, ventricle at 15 dpc, AV canal at 15 dpc. For a detailed description of the library from which this clone was derived, please visit our web site at ratest.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996) TAG_LIB=U1-R-BJ2		

[illegible]

RESULT		6		
BFL120840				
LOCUS				
DEFINITION	BFL120840	823 bp	mRNA	EST
	601757442n1	NCT_CGAP_Mam3 Mus musculus cDNA clone IMAGE:		24-OCT-2000 3986651 5'
ACCESSION	BFL120840			
VERSION	BFL120840..1	GI:10959880		
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
REFERENCE	Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murinae; Mus.			
AUTHORS	1 (bases 1 to 823)			
TITLE	NIH-MGC http://mgc.nci.nih.gov/.			
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT	Unpublished (1999)			
Contact:	Robert Strausberg, Ph.D.			

FEATURES	Location/Qualifiers
source	1. .823

BASE COUNT ORIGIN	150 a	226 c	257 g	190 t
----------------------	-------	-------	-------	-------

Query Match	30.1%	Score 312.2	DB 11	Length 823
Best Local Similarity	76.5%	Pred. No. 2.1e-33		
Matches 423	Conservative	0	Mismatches 123	Indels 7
				Gaps 3

QY	55	ctgagccacatctccgaagaagccctgattgttcacacaggagcctggcctggcctgta	114
Db	186	CTGGTGGCCATCGCGGCGAGCCTGTGGATTGCCAAGTGGGAGATGGTCTGTGGCTGGCTGA	245
QY	115	gqcaagcagtgccctctcaatgcgctgggttttgcagtagcctggagaccatccctcgtgc	174
Db	246	GGCAGCATGTCGCTCTTCATTGCTGGCTTGGCGGTGGGCTGGGGAGCCATCCCTGGCTGC	305
QY	175	ctcatgcaagaagatcttcctcctcctgcacatcaagaggtgtggtctacaggcgcctctgtccctc	234
Db	306	CTCATGTAGAGATTTTCCCTCTGCATGTCAAGGGTGTGGCTACCGGCACTCTGTGTCTC	365
QY	235	accacactgattcagcctcttctgtgtgaccaaagatttcaacagacatcaltggaatcctc	294
Db	366	ACCACTGGTTCATGGCTTCTTAGGACCAAGAGTTCAGCAGCGTCATGGAGATCCTC	425
QY	295	agaccctagcgagcctctcgtctaacacgtctgctctctctatcctcaagctccttctcaag	354
Db	426	AGACCTTAGCGGCTTCTCTGGCTCAACCGCTGTCTTGTGCTCTCACTGTCTCTTTGGCAC	485

QY	355	ctcacccttgcctcttgagactaaagcagagactcttgaaacaatc-aaacccactttgag	413
Db	486	ACTGACCCTGTCTCTTGAGACTAAAGGACGAGCTGTGGACACATGCGACAGCCATTTCGA	545
QY	414	ggacagctgacagcagaccctctctgtgactlgcagccctgagctgagctgctcgatttc	473
Db	546	GCGGCGATGTGACAGACCCCTTTCTGTG-CTGGAGAGCCGCCGACTGTGAGCTGGCTTGGGGTTTC	604
QY	474	aaagagagtgagatgagc-----ctcagctgagccagagtttgagcccgagggccccctgact	528
Db	605	AGAGGAGATGTGAGTGGCTTGAACCTAAAGGCACAGCTCAGGTGAGAGCTTGGAGGCCCTTAAT	666
QY	529	cctcaagaatctccggccagactctgtccagactctcaaccagagatctccacacatgagctc	588
Db	665	CCTCACTGAAGAGCGCTCTTTGGCGACAACTTTAAAGCCAGATTCCACCGGTGAAGCGC	724
QY	589	accagatctctgag	601
Db	725	AGCGGAATCTGAG	737

RESULT		7							
BG749509									
LOCUS	BG749509	949 bp	mRNA	EST	15-MAY-2001				
DEFINITION	602070766.f1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4844411 5'								
	mRNA sequence.								
ACCESSION	BG749509								
VERSION	BG749509.1	GI:14060162							
KEYWORDS	EST.								
SOURCE	human.								
ORGANISM	Homo sapiens								
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								

Plate: LLCM1681 row: p column: 12
High quality sequence stop: 745.

FEATURES	Location/Qualifiers
source	1. .949

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:484441"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site.1: XhoI; Site.2:
EcoRI; cDNA made by oligo-dT priming. Additionally,
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)
>Note: This is a NIH_MGC Library. |"
BASE COUNT
160 a 302 c 264 g 223 t

```

Query Match	25.1%	Score 271	DB 11	Length 949
Best Local Similarity	84.7%	Pred. No. 6	9e-28	
Matches 304	Conservative 0	Mismatches 55	Indels 0	Gaps 0
57	ggatccatctcgcgaagacgcbtgcgttcaacttggctggtcgtgcgtgtagg	116		

Db 64 GGGCTGTCTCTGTGACAGGCTGTTGATGCCAGCTGGGGCTGGCTGGCTGGCTGGG 123
 QY 117 cagactgtcctctcaccagctgtttgagtagagctgggagccatccctgtcct 176
 Db 124 CAGATGTGCTCTTCATCAGCGCGGCTTTGGCGTGGGCTGGGAGCCATCCCTGGCTCT 183
 QY 177 catgtcagaatcttccctgtacatcaagaagtgtgtgtacacggcgtctgtgtcctac 236
 Db 184 CATGTCAAGATCTTCCCTCTGTGATGTCAGAGGCGGTGGCGACAGCATCTCGTCCCTAC 243
 QY 237 caactgttcaatgacctcttggtgaccagaagagtttaacagcatatggagatcctcag 296
 Db 244 CAATGGCTCATGGCTTTCTTCGTGACCAAGAGGTTCAGCAGCTCATGAGGTCCTCAG 303
 QY 297 accctagcgagctcttgctcaccgctcctgtatcctcagagctccttcaagct 356
 Db 304 GCCCTATGAGAGCTTTGGGCTTGGCTCCGCTTCTGCATCTTCAGTGTCTTTTACCTTT 363
 QY 357 cactttgtccctgagactaaagcaggaactctggaacaatcacagccatltgaagg 415
 Db 364 GTTCTGTGCTCCCTGAAGACTTAAGGAAGACTGTGAGACAAATCAGCCCATTTTGAAG 422

RESULT 8
 BE231636 546 bp mRNA EST 10-JUL-2000
 LOCUS BE231636 136437 MARC 1Pig Sus scrofa cDNA 5', mRNA sequence.
 DEFINITION BE231636
 ACCESSION BE231636
 VERSION BE231636.1 GI:9016354
 KEYWORDS EST.
 SOURCE pig.
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 546)
 Fahnenkrug,S.C., Freking,B.A., Rohner,G.A., Smith,T.P.L., Casas,E.,
 Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
 and Keele,J.W.
 Design and use of two pooled tissue normalized cDNA libraries for
 EST discovery in swine
 Unpublished (2000)
 Contact: Smith rpl
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt-trimmed with phred
 v0.980904.e. Vector identified by cross-match with the -m1nscore 18
 and -mismatch 12 options.
 PCR primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCACAGTACAGCAGC
 Plate: 70 row: A column: 21
 Seq primer: ATTTAGGTGACATCTATG.
 FEATURES
 source Location/Qualifiers
 1..546
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="MARC 1Pig"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORF6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from day 11, 13, 15, 20,
 and 30 embryos."

BASE COUNT 76 a 182 c 162 g 124 t 2 others
 ORIGIN

Query Match 25.9%, Score 268.2; DB 10; Length 546;
 Best Local Similarity 82.0%; Pred. No. 2e-27;
 Matches 309; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 73 gagcgtctgattgtcaacctgaggctgctgctgctgtatgaagacatgtgccttc 132
 Db 2 GAGCCCGCCATGACCACGTGGGCTGGGCTGGCTGGCGGTGGGCGACAGTGTGCTCTTC 61
 QY 133 atcgtgttctgagtagagctgggagccatccctcggctcctcctcgttaagaatcttc 192
 Db 62 ATTGCTGTGCTGGCGCTGGGCTGGGAGCCATCCCTGGCTCTCATAGTCTTGAAATCTTC 121
 QY 193 cctctcacatcaagggtgtgtgtacccggtctgtgtcctcaaccaactgttcatggcc 252
 Db 122 CCTCTGCACATCAAGGCGCGTGGCCACCGCGCTGTGTCTCTCAACCACTGGCATATGGCC 181
 QY 253 ttctgtgagccaaagggttaacagataatgagatcctcagaccctcagcgcttc 312
 Db 182 TTTCGTGTGACAAAGAGTTTCAGCAGCTCATGAGGTCTCAGGCCCTTACGGGTCTTC 241
 QY 313 tgcctacagctgctctgtatcctcagcgtccttcaacgtccttgccttgag 372
 Db 242 TGGCTGCGCTCCGCGCTTCTGCATCTTCAGTGTCTTTTCACTGTGGCTGTGCTTGA 301
 QY 373 actaaagcagagactctggaacaatcacagccatltgaggagcgtgacgagccctt 432
 Db 302 ACCAAGGGAAGACTGTGAGCAGATCACGCCATTTTGAAGGCGGAGTACAGCCTTC 361
 QY 433 tctgtactgtgagccc 449
 Db 362 CCCGAGAGTGGCCACCC 378

RESULT 9
 BF420257/c 295 bp mRNA EST 28-NOV-2000
 LOCUS BF420257 UI-R-BJ2-dpk-g-03-0-01.s1 UI-R-BJ2 Rattus norvegicus cDNA clone
 DEFINITION UI-R-BJ2-dpk-g-03-0-01 3', mRNA sequence.
 ACCESSION BF420257
 VERSION BF420257.1 GI:11408246
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 295)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msocares@blue.weeg.uiowa.edu
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to identify it as a clone from the
 normalized ventricle at 16.5 dpc library cDNA library Preparation:
 M.B. Soares Lab Clone distribution: clones will be available
 through Research Genetics (www.resgen.com) The following repetitive
 elements were found in this cDNA sequence: 1-21,
 >AT-rich#low-complexity
 Seq primer: M13 Forward
 POLYA=Yes.
 FEATURES
 source Location/Qualifiers
 1..295
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"

```

/clone="UI-R-BJ2-bpk-g-03-0-0-UI
/clone_lib="UI-R-BJ2"
/lab_host="DH10B (Life Technologies)"
/notes="vector: pRT73b-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; The UI-R-BJ2
library is a subcloned library derived from the following
tissues: heart, atrium at 15 dpc, ventricle at 16.5 dpc,
atrium at 16.5 dpc, ventricle at 13 dpc, ventricle at 15
dpc, AV canal at 15 dpc. For a detailed description of
the library from which this clone was derived, please
visit our web site at ratseq.iglow.edu. The
subtraction has been previously described in (Donaldo,
Lennon and Soares, Genome Research 6:791-805, 1996)
TAG_LIB=UI-R-BJ2
TAG_TISSUE=ventricle at 16.5 dpc
TAG_SEQ=GTTCG
TAG_SEQ=63 c 68 g 88 t

```

Query Match	23.9%	Score 248.2	DB 11	Length 295
Best Local Similarity	96.3%	Pred. No. 13e-24		
Matches 286	Conservative	0	Mismatches 8	Indels 3
				Gaps 3

[illegible]

Qy 764 ggcgaagtgtaagtcacgagaagaataatgtctctgtctgtcacaaccaagluccctctc 823
|||||
|||||
Db 235 ggTCAGGTGATCCAGAGAAGAAAGAGAT - GTTCTTGCTCTTGTCACCAACCAAGTCCTTCTC 177

Qy 824 agagtgc-ccaaagacgtccgattcacccttggttagccagcttaccatcaactaacag 882
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 176 AGAGTGGCCGAGAGACCTCGGATTACACTTTGGGTTAGCCAGCCTTAACCAATCACCTTACAG 117

QY 883 gttctctccaactcttcagctgtgtctcaagtgltccfgatcatctagtcaccaggtctgttg 942
|||||
116 GTTCTCTCAACTCTTTCCTTGCTCTCAAGTGTCTCGATGATATTAGTCACCAAGCTTTTG 57
Db

QY 943 agttccagaaataaaagcctcttcgcgttcaaaaaaaaaaaaaaaaaaaaaa 999

Db 56 AGTTTCAG-AAAATAAAGGCGCTTTCCTCACTTTAAAAA 1

RESULT 10
R551290

LOCUS	BF551290	426 bp	mRNA	EST	12-DEC-2000
DEFINITION	UI-R-C0-bq-f-01-0-UI.r1 UI-R-C0 Rattus norvegicus cDNA clone				
UI-R-C0-bq-f-01-0-UI 5'	, mRNA sequence.				
ACCION	BF551290				

VERSION	Bf551290.1	GI:11661020
KEYWORDS	EST.	
SOURCE	Norway rat.	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

AUTHORS	Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery

COMMENT
MEDLINE
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and

Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel.: 319 335 8250
Fax: 319 335 9565
Email: mscoates@blue.weeg.uiowa.edu
CDNA Library Preparation: M.B. Soares Lab Clone Distribution:
clones will be available through Research Genetics (www.resgen.com)

This clone is also available through the I.M.A.G.E. Consortium at LNL (info@image.llnl.gov). IMAGE ID= 1773689
Seq primer: M13 Forward.

FEATURES	Location/Qualifiers
source	1. .426

```

/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CO-hq-f-0-0-UI"
/clone_lib="UI-R-CO"
/dov_staco="adult"

```

/lab_host="BH10B/Life Technologies)"
 /note=Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker: Site_1: Not I; Site_2: Eco RI; The UI-R-C0
 library is a subcloned library derived from the UI-R-A1
 and UI-R-E1 libraries. The UI-R-A1 library consisted of a
 mixture of individually tagged normalized libraries
 constructed from rat placenta, adult lung, brain, liver,
 kidney, heart, spleen, ovary, and muscle. The UI-R-E1
 library consisted of a mixture of individually tagged
 normalized libraries constructed from 8-, 12 and 18-day
 embryo. The tag is a string of 3-5 nucleotides present
 between the Not I site and the oligo-dT track which
 allows identification of the library of origin of a clone
 within the mixture. The subcloned library (UI-R-C0) was
 constructed as follows: PCR amplified cDNA inserts from a
 pool of UI-R-A1 and UI-R-E1 clones from which 3' ESTs had
 been derived was used as a driver in a hybridization with
 the pooled UI-R-A1 and UI-R-E1 library in the form of
 single-stranded circles. The remaining single-stranded
 circles (subtracted library) was purified by
 hydroxyapatite column chromatography, converted to
 double-stranded circles and electroporated into BH10B
 bacteria (Life Technologies) to generate the UI-R-C0
 library. This procedure has been previously described
 (Bonaldo, Lennon and Soares, Genome Research 6: 791-806,
 1996)

BASE COUNT	99 a	129 c	105 g	93 t
ORIGIN				

Query Match	22.38;	Score 231.6;	DB 11;	Length 426;
Best Local Similarity	82.58;	Pred. No. 1.9e-22;		
Matches 354;	Conservative	0;	Mismatches 5;	Indels 70;
				Gaps 5;

Qy 511 ccacagggccccctgactcctcagatattccgggacagcttgtccagatcacaaccaga 570
|||||
Db 1 ccacagggccccctgactcctcagatattccgggacagcttgtccagatcacaaccaga 60

QY 571 ttcacacacatgagcttcaccagattctgagctctntgagccctgctgcacacacagcac 630
|||||
Db 61 -ttcacacacatgagcttcacacagattctgagctctntgagccctgctgcacacacacac 118
|||||

QY 631 att----- 633
||
Db 119 ATCTGGGGAGCCATGCTTCGCCCTAACCCACGGGTCCTCTCTTCCTAAGACTAAGCAAC 178

[illegible][illegible]

745 gatttggctcagaacaaggtcagtgtagtccaggaagaagaatggtctctct 804

805 tgcacaccaaagtcctctcagagtg-ccaaagactccgattcaccttgggttagcca 863

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

Oy 864 gctaccoca 872
Db 418 GCTTACCCA 426

RESULT 11
LOCUS BE910478
DEFINITION 601501005F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3903048 5', mRNA sequence.
ACCESSION BE910478
VERSION BE910478.1 GI:10407108
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 934)
NIH-MGC <http://mgs.nci.nih.gov/>
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LAM9706 row: 1 column: 01
High quality sequence stop: 710.

FEATURES
source
1..934
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3903048"
/clone_lib="NIH_MGC_70"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.1 Kb. Library constructed by Life Technologies."

BASE COUNT 143 a 291 c 287 g 213 t
ORIGIN

Query Match 22.3%; Score 231.6; DB 11; Length 934;
Best Local Similarity 80.6%; Pred. No. 1.4e-22;
Matches 295; Conservative 0; Mismatches 69; Indels 2; Gaps 2;

Oy 57 ggtgccatcttcgcagagcctgcgtgattcaactggagctgctgctgctgctgtag 116
Db 359 GCGGGCTGCTCTGCACACCGCTGTGATGCCAGCGGGGCTGCGCTGCGCTGGG 418

Oy 117 cagcactgacctcctcctcctgctgctgctgctgctgctgctgctgctgctgct 176
Db 419 CAGCATGCGCTCTTCATCGCCGCTTGGCGGTGGGCGCCATCCCTGCTGCTCT 478

Oy 177 catgctagagatctcctctcgtacatcaagggtgtgctcagcgctgtgtctcag 236
Db 479 CATGTCAAGATCTTCCTCTGCATGTCAGAGCGCTGGGAGACAGCATCTGCTCTCAG 538

Oy 237 caactggtctatgctctctctgtgtgaccaaagatttaacagcatcagagatccctcag 296
Db 539 CAACGTGCTCATGGCTTCTCTGACCAAGAGATTACAGACCTCATGAGAGTCTCTAG 598

Oy 297 accctacggcgctctcgtgctcagcgctgctctgtatcctcagcgtccttlltaagct 356
Db 599 GCCCATGAGACCTTCTGCTTGGCTTC-CCTTCTGATCTTCACTGATGCTCTTTTCACT 657

Oy 357 caacttgcctctagactaaagcagagactctgaacaatcacagccactlltgagga 416

Db 658 TTGTTCTGCTGCTGACACATGAGGAAGACTCTGAGAC-AATCACACCATTTTGAGGG 716
Oy 417 cagtgca 422
Db 717 CGATGA 722

RESULT 12
LOCUS BG290178
DEFINITION 602385265F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4514200 5', mRNA sequence.
ACCESSION BG290178
VERSION BG290178.1 GI:13046711
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 712)
NIH-MGC <http://mgs.nci.nih.gov/>
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LAM10402 row: a column: 17
High quality sequence stop: 707.

FEATURES
source
1..712
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4514200"
/clone_lib="NIH_MGC_93"
/tissue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. oligo-dT primed. Average insert size 1.7 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."

BASE COUNT 117 a 221 c 211 g 163 t
ORIGIN

Query Match 22.3%; Score 231.4; DB 11; Length 712;
Best Local Similarity 85.6%; Pred. No. 1.6e-22;
Matches 292; Conservative 0; Mismatches 46; Indels 3; Gaps 3;

Oy 90 cctgggagctgagctgctgctgctgctgctgctgctgctgctgctgctgctgct 149
Db 5 CGTGGGCTGGGCTGCTGCGCGCGGAGCATGTGCTCTTCATCGGCGG-CTTGGCGGT 63

Oy 150 agctggggagccatccctctgctcctcatgtcagaagatctcctctgcacatcaagg 209
Db 64 GGGCTGGGGGCCCATCCCTGCTCTCATGTCAAGATCTTCCTCTGATGTAACGG 123

Oy 210 tgtgctacccggcgctctgtctcctcaactggtctatgctcttctcgtgacccaaga 269
Db 124 CGTGGCGACAGGATCTGCTGCTACCAACTGCTCATGCGCTTCTCGTAGCCAAAGA 183

Oy 270 gtttaacagatc-atgagatcctcagacacctcagcgctctcgtgctcagcgtgct 328
Db 184 GTTCAGCAGCGCTCATGAGAGTCTCTCAGCGCTCATGAGACCTTCTGCTTCCCTT 243

Oy 329 tctgtatcctcagcgctcctlltcaagctacacttgtctcctgagactaaagcagagact 388

Db	244	TCTGATCTCTGACGTGTCCTTTTCAC-TTGTTCTGTGTGTCCTTAACCTAAGCAAGACATC	302
OY	369	tggaacaatcacaagaccatttgaggacagtgacgaccc	429
Db	303	TGGAACAATCATCAGCCCATTTGAGCGCGGATGACACCCAC	343
RESULT 13			
LOCUS	H34372	324 bp	mRNA
DEFINITION	EST111272 Rat PC-12 cells, NCF-treated (9 days)	EST	02-APR-1998
ACCESSION	clone RPNT28	3' end, mRNA sequence.	Rattus sp. cDNA
VERSION	H34372		
KEYWORDS	H34372.1	GI:979789	
SOURCE	EST.		
ORGANISM	Rattus sp.		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1 (bases 1 to 324)		
AUTHORS	Lee,N.H., Weinstein,K.G., Kirkness,E.F., Earle-Hughes,J.A., Fuldner,		
	R.A., Mammar,S., Glodek,A., Gocayne,J.D., Adams,M.D., Kerlavage		
	,A.R., Fraser,C.M. and Venter,J.C.		
	Comparative expressed sequence-tag analysis of differential gene		
	expression profiles in PC-12 cells before and after nerve growth		
	factor treatment		
	Proc. Natl. Acad. Sci. U.S.A. 92, 8303-8307 (1995)		
TITLE	95396786		
JOURNAL	Contact: Lee, NH		
MEDLINE	The Institute for Genomic Research		
COMMENT	9712, Medical Center Drive, Rockville, MD 20850, USA		
	Tel: (301)-838-3529		
	Fax: (301)-838-0208		
	Email: nhlee@tigr.org		
	For clone availability please contact the TIGR database		
	(tcbinfo@tigr.org)		
	Seq primer: M13 - 21.		
FEATURES	Location/Qualifiers		
source	1..324		
	/organism="Rattus sp."		
	/db_xref="ATCC (inhost):2005377"		
	/db_xref="taxon:10118"		
	/clone="RPNT28"		
	/clone_1lb="Rat PC-12 cells, NCF-treated (9 days)"		
	/note="Vector: pBluescript SK-; Site 1: EcorI; Site 2:		
	XhoI; poly(A)+ RNA was purified from 9-day NCF treated		
	PC12 cells. cDNA was constructed using an oligo-dT primer		
	and directionally cloned using the Lambda Zap II Vector		
	kit by Stratagene"		
BASE COUNT	89 a 74 c 79 g 78 t 4 others		
ORIGIN			
Query Match	22.1%:	Score 229:	DB 11: Length 324:
Best Local Similarity	95.5%:	Pred: No. 4.7e-22:	
Matches 298:	Conservative 0:	Mismatches 8:	Indels 6: Gaps 6:
OY	670	tttgagggtgtgtgtcctcaagcaac-tgccatacc-t-aacttgactggggaatgaga-	726
Db	314	TTTTGGGCGTCTTGAGCACTTAAGCAACTTGGCCCTTACTTCACTTGAGCTGGCGCATANAAA	255
OY	727	agggaactagccacataaatttggtctcaagaacaagtgagtgagtcaggagaagaa	786
Db	254	AGGCACTTAAGCCACATTAACATTTGGGCTTCAGAAACAAGCTCAGGTGATGCCAGAGAA	195
OY	787	agagaatggctcttgctgtglaacaagaatcccttcacagatg-ccaagaacctcogat	845
Db	194	ACAGAAAT-GTTCTTGCTGTGTCACCAAGTCTTCTCAGAGTGGCCAGAGACTCCGAT	136
OY	846	tcaacttggggttagccagctaccatcaacttacaaggtctctcccaactcagctgat	905

Db	135	TCACCTTGGGGTTACCAAGCTTACCCAINNACTTACAGGTTCTCTCCACACTCTTCCCTTGT	76
QY	906	ctcaagtcctcgtatcatatagtcaccaggctctgttgagttcagaatacaaaagcct	965
Db	75	CTCAGTGTCTCTGGATCATATTAGTCACAGGCTCTTGTGAGTTTCAG-AAAATNMAAGGCT	17
QY	966	cttcctcctcaaa 977	
Db	16	CTTTCCTGTTCAA 5	
RESULT	14		
LOCUS	AL565390	752 bp	mRNA
DEFINITION	AL565390 LTI_FL013_FBrn1	Homo sapiens	CDNA clone CS0DF005YE02 3
ACCESSION	AL565390		prime, mRNA sequence.
VERSION	AL565390.1	GI:12916718	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 752)		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	unpublished (2001)		
FEATURES	Contact: Genoscope		
source	Genoscope - Centre National de Sequencage		
	BP 191 91006 Evry cedex - France.		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
	Location/Qualifiers		
	1..752		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0DF005YE02"		
	/clone_lib="LTI_FL013_FBrn1"		
	/dev_stage="pooled tissue from post conception fetuses (20		
	week, 24 week and 26 week)"		
	/lab_host="DH10B"		
	/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand		
	CDNA was primed with a NotI-colligodf) primer. Five prime		
	end enriched, double-stranded CDNA was digested with Not I		
	and cloned into the Not I and Eco RV sites of the		
	pCMVSPORT 6 vector. Library was constructed by Life		
	Technologies. Contact : Feng Liang Life Technologies, a		
	division of Invitrogen 9800 Medical Center Drive Rockville		
	, Maryland 20850, USA Fax : (1) 301 610 8371 Email :		
	fliang@litech.com URL :		
	http://fulllength.invitrogen.com"		
BASE COUNT	177 a	206 c	242 g
ORIGIN		118 t	9 others
Query Match	21.1%	Score 218.6:	DB 10; Length 752;
Best Local Similarity	79.5%	Pred. No. 8.4e-21:	
Matches	290;	Conservative	2; Mismatches 67; Indels 6; Gaps 3;
QY	57	gtatgccaatcttcggaagagcgtcgtcgtatcactcttgaggctcgtcgtcgtcgtatg	116
Db	731	SAGGCGCTGTCTTGACAGCGCTGTGATCCAGCGCTGGGCTSGCTCGCTGCGN	672
QY	117	cagagatgagcctctcat--cgctcgatttcagtagccttgaggaccatcccc-tggc	172
Db	671	CANCATGTGCGCTTTTCNATCGCCACGCTTTTGCGGTGGGCTGGGGCCCATGCCCTTGGC	612
QY	173	tcctcatgacagagatctt--ccctctgacacatcaagggtgtgagctaccgagctgtgt	230
Db	611	TGCTCATGTGACAGATCTTCCCTCCCTCGCATCTTTCACAGGCGCGACAGAGCATCTGCGT	552
QY	231	cctaaccaactggtgtcatatgagccttctgtgtgacccaagaagtttaacagatcatatg	290
Db	551	CCNACCAAGATGCGTCATATGGCTTTCTGTGTGACCAAGATGTCACACGATCTATAGAGGT	492

```

OY 291 cctcaagaccctacgagcgtctgtgtcaccgctgcttctgtatcctcaggtcctt 350
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 491 CCTCAGCCCTATGAGAGCTTGTGGCTTGGCTTGTGATGTGAGTGTCTTTT 432
OY 351 cagcgtcacccttgtgtccttgaagcagcagcgtctgtgaacaaatcacagccatt 410
      ||| | | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 431 CACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 372
OY 411 gaagg 415
      ||
Db 371 TGAGG 367

RESULT 15
A2985752 796 bp DNA GSS 27-APR-2001
LOCUS 2M026701.R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
DEFINITION clone UUGC2M026701.R, DNA sequence.
ACCESSION A2985752
VERSION A2985752.1 GI:13856979
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 796)
Dunn,D., Aoyagi,A., Barber,M., Beacom,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Rellily
,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0267 row: 0 column: 11
Seq primer: CACACAGGAACACAGCTATGAC
Class: plasmid ends
High quality sequence stop: 796.
Location/Qualifiers
1..796
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M026701"
/clone_1lb="Mouse 10kb plasmid UUGC2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv. Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into

```

```

BASE COUNT 171 a 217 c 248 g 160 t
ORIGIN
Query Match 19.2%; Score 199.6; DB 13; Length 796;
Best Local Similarity 79.2%; Pred. No. 36-18;
Matches 290; Conservative 0; Mismatches 61; Indels 15; Gaps 4;
OY 285 ggaatcctcgaagaccctacgagcgtctgtgtcaccgctgcttctgtatcctcaggt 344
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 490 GGAGATGCTCAGACCTACGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 431
OY 345 ccttccagctacacttgtctccttgaagcagcagcgtctgaacaaatcacagc 404
      ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 430 CTTATTTCACACTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 371
OY 405 ccatlgaaggagcgtgagcagcccttctgtgactgagcagcctgagctgagctgct 464
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 370 CCATTTGAGAGGACGATGACAGACCCCTTCTGTGTGTGTGTGTGTGTGTGTGTGT 311
OY 465 tggggttcaaaaggagtgagtgagct-----cagtgaccacagtttgagccag 516
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 310 TTGGGTTTCAAGAGAGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAG 251
OY 517 gggccctgactcctcagatltccgggcaagcttgttcagatcctaacccagattccac 576
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 250 GGCCCTTGACTCTC--ACCTAGAGGCCCTTTTGGCCAGATCTCAGCCCAATTCC-C 194
OY 577 accatgagcttcacagatctcga---ggctcctnagcctcgtcgaacacagacat 632
      ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 193 ACCGTGAGGCCACACAGACTGTGAGTGCGGTCTGTGACAGCCTCTTCACACAGGACAT 134
OY 633 ttgagg 638
      |||
Db 133 CTGAGG 128

```

Search completed: February 13, 2002, 13:37:40
Job time: 2987 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 15:02:23 Search time 4552.73 Seconds
(without alignments)
1021.849 Million cell updates/sec

Title: US-09-516-493-11

Sequence: 1 gagcctgtgatgttcaact.....tcclcagcgtcccltllcacy 282

Scoring table: IDENTITY_NUC

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Pre-processing:	Minimum Match	0%
Post-processing:	Minimum Match	0%

Listing first 45 summaries

Database :

1. gm_ba: *
2. gm_htg: *
3. gm_in: *
4. gm_om: *
5. gm_ov: *
6. gm_pat: *
7. gm_ph: *
8. gm_pl: *
9. gm_pr: *
10. gm_ro: *
11. gm_srs: *
12. gm_sv: *
13. gm_un: *
14. gm_v1: *
15. gm_ba: *
16. gm_fun: *
17. gm_hum: *
18. gm_in: *
19. gm_om: *
20. gm_ot: *
21. gm_ov: *
22. gm_pat: *
23. gm_ph: *
24. gm_pl: *
25. gm_ro: *
26. gm_srs: *
27. gm_sv: *
28. gm_un: *
29. gm_v1: *
30. gm_hgo_hum: *
31. gm_hgo_in: *
32. gm_hgo_inv: *
33. gm_hgo_rod: *
34. gm_hg_hum: *
35. gm_hg_inv: *
36. gm_hg_rod: *
37. gm_hg_other: *

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	282	100.0	2087	6	AX076669	AX076669 Sequence
2	282	100.0	2087	10	RNO245935	AJ245935 Rattus norvegicus
3	282	100.0	2189	10	AB033418	AB033418 Rattus norvegicus
4	242.6	86.0	2072	6	AX076671	AX076671 Sequence
5	242.6	86.0	2072	10	MM0245936	AJ245936 Mus musculus
6	241	85.5	1490	10	MM017802	Y17802 Mus musculus
7	241	85.5	1843	10	AF232061	AF232061 Mus musculus
8	220.2	78.1	1012	4	AF321324	AF321324 Bos taurus
9	219.2	77.7	1012	4	AF321324	Y17801 Homo sapiens
10	219.2	77.7	2080	6	AX119740	AX119740 Sequence
11	217.6	77.2	1873	6	HS245937	AJ245937 Homo sapiens
12	217.6	77.2	2217	6	AX076667	AX076667 Sequence
13	125.2	44.4	225370	9	AL445222	AL445222 Human DNA
14	89.4	31.7	1580	6	AX072418	Y17808 Homo sapiens
15	89.4	31.7	2011	6	AX076679	AX076679 Sequence
16	89.4	31.7	2487	2	AC010025	AC010025 Drosophila
17	89.4	31.7	2487	2	AC010025	AC010025 Drosophila
18	88.8	31.5	42845	2	AC010025	AE003340 Drosophila
19	88.8	31.5	185413	2	AE003340	Y076677 Sequence
20	88.8	31.5	297385	3	AE003340	AC001644 Genomic s
21	87.8	31.1	1541	6	AX076677	AC001644 Genomic s
22	71	25.2	42415	9	AC002325	AC002325 Homo sapiens
23	71	25.2	44762	9	AC002325	AC002325 Homo sapiens
24	65.4	23.2	127914	2	AC083945	AC083945 Oryza sativa
25	65.4	23.2	147706	8	AC083945	AL161544 Arabidopsis
26	60.4	21.4	198735	8	ATCC46	AC091762 Mus musculus
27	60.4	21.4	213788	8	ATCC46	AC091762 Mus musculus
28	60	21.3	203354	2	AC091762	AC091762 Mus musculus
29	59.4	21.1	2533	4	BOVSLUT1	AC008004 Drosophila
30	59.2	21.0	175506	3	AC008004	ME5072 Leishmania
31	58.6	20.8	2798	3	LEIDITFA	M21747 Rabbitt giuc
32	57.8	20.5	2487	4	RABEGT	AC019887 Drosophila
33	56.2	19.9	44543	2	AC019887	AC011252 Drosophila
34	56.2	19.9	157966	3	AC011252	AC008311 Drosophila
35	56.2	19.9	161117	3	AC008311	AE003601 Drosophila
36	56.2	19.9	303626	3	AE003601	X75440 C. kessleri
37	55.4	19.6	1605	8	CKRNAHP3	AJ278765 Lycopersicon
38	55.2	19.6	3399	8	LES278765	AC017643 Drosophila
39	55.2	19.6	177818	2	AC017643	AJ24968 Arabidopsis
40	55	19.5	1578	8	ATH24968	X17058 Sus scrofa
41	53.4	18.0	2230	4	SSGIJUP	AF367260 Arabidopsis
42	53.4	18.9	1729	8	AF367260	S77924 glut-1-glucose
43	53	18.8	321	10	S77924	M22998 Mouse facili
44	53	18.8	2473	10	MUSGLITRN	AC016794 Drosophila
45	53	18.8	63855	2	AC016794	

CDS

/db_xref="taxon:10118"
 30..1466
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC28496.1"
 /db_xref="GI:12711201"
 /translation="MSPEDPQETOPFLRSPGARAPGRVFLATPAALGPLSPFAL
 GYSSPAISPLRTAPALRLGDTASMGAVVTLGAAGVGLGMLDRAGRLSLT
 CVPPTGFAVITTAARDVMMILGGRLTGLAGVASLVAPLYISLAVPAVRLGSC
 VOLMYVTGILLAVAGVLEMMILAVGCPPTMLLMCYMPETPRFLTHOYOA
 MAALFPLWSEBGEWEPVGAHQFOALMRPVGHPPLIGICIMVFOOLSGVNAI
 MRYANTIEEAKFKDSSLASVYTGIIQVLFPAVALIMDRAGRLILASGVIMFPM
 SAFGYFKITQSGPSSSHVGLVPIISAPADVHLGLAMLAAGSMCLFIAGFAVGMP
 IRLMLSELPPLHIKGVATGCVILNMFMAFLVTKERNIMEILRPGAFWLTAAFCI
 LSVLETFPEPTKGRTELOITAHFEGR"
 BASE COUNT 355 a 673 c 580 g 479 t
 ORIGIN

Query Match 100.0%; Score 282; DB 6; Length 2087;
 Best Local Similarity 100.0%; Pred. No. 1.2e-59;
 Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaagctgctgtagttaccctggcgctgctgctgctgctgctgctgctc 60
 |||
 DB 1116 GAGCCTGCTGANTGACCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTTC 1175
 |||
 QY 61 atcgctgcttgcagtagtgcgagccatccctgctcctcctcagagatcttc 120
 |||
 DB 1176 ATCGCTGCTTTCAGTAGAGCTGGGACCCATCCCTGGCTCTCATGTCAGAGATCTTC 1235
 |||
 QY 121 cctctgacatcaagggtgtgctaccgctgctgtctcctacacactgctcagcc 180
 |||
 DB 1236 CCTCGACATCAAGGGGTGTGCTACCGGCTGTGTGCTCACCACAGCTTCATGAGCC 1295
 |||
 QY 181 ttctgtgtgaccaaagagtttaacagacatcagagatcctcagagccctc 240
 |||
 DB 1296 TTCTGTGTGACCAAGAGTTTAAACAGCATCATGAGATCCTCAGAGCCCTTACGCGCTTC 1355
 |||
 QY 241 tggctaccgctgctctctgatactcagagctccttcaag 282
 |||
 DB 1356 TGGCTACCGCTGCTCTGTATCTCAGCGCTCTTTTCACG 1397

RESULT 2

LOCUS RNO245935 2087 bp mRNA ROD 18-FEB-2000
 DEFINITION Rattus norvegicus mRNA for glucose transporter (GLUTX1 gene).
 ACCESSION AJ245935
 VERSION AJ245935.1 GI:7018604
 KEYWORDS glucose transporter; GLUTX1 gene.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 2087)
 AUTHORS Ibberson M., Oldry M. and Thorens B.
 TITLE GLUTX1, a novel mammalian glucose transporter expressed in the
 central nervous system and insulin-sensitive tissues
 J. Biol. Chem. 275 (7), 4607-4612 (2000)
 JOURNAL 20138191
 MEDLINE 2 (bases 1 to 2087)
 REFERENCE Ibberson M.R.
 AUTHORS Direct Submission
 TITLE Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
 and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
 CH-1005, SWITZERLAND

FEATURES

source 1..2087
 /organism="Rattus norvegicus"
 /db_xref="taxon:10118"
 30..1466
 gene

CDS

/gene="GLUTX1"
 30..1466
 /gene="GLUTX1"
 /codon_start=1
 /product="glucose transporter"
 /protein_id="CAB75729.1"
 /db_xref="GI:7018605"
 /translation="MSPEDPQETOPFLRSPGARAPGRVFLATPAALGPLSPFAL
 GYSSPAISPLRTAPALRLGDTASMGAVVTLGAAGVGLGMLDRAGRLSLT
 CVPPTGFAVITTAARDVMMILGGRLTGLAGVASLVAPLYISLAVPAVRLGSC
 VOLMYVTGILLAVAGVLEMMILAVGCPPTMLLMCYMPETPRFLTHOYOA
 MAALFPLWSEBGEWEPVGAHQFOALMRPVGHPPLIGICIMVFOOLSGVNAI
 MRYANTIEEAKFKDSSLASVYTGIIQVLFPAVALIMDRAGRLILASGVIMFPM
 SAFGYFKITQSGPSSSHVGLVPIISAPADVHLGLAMLAAGSMCLFIAGFAVGMP
 IRLMLSELPPLHIKGVATGCVILNMFMAFLVTKERNIMEILRPGAFWLTAAFCI
 LSVLETFPEPTKGRTELOITAHFEGR"
 BASE COUNT 355 a 673 c 580 g 479 t
 ORIGIN

Query Match 100.0%; Score 282; DB 10; Length 2087;
 Best Local Similarity 100.0%; Pred. No. 1.2e-59;
 Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaagctgctgtagttaccctggcgctgctgctgctgctgctgctgctc 60
 |||
 DB 1116 GAGCCTGCTGANTGACCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTTC 1175
 |||
 QY 61 atcgctgcttgcagtagtgcgagccatccctgctcctcctcagagatcttc 120
 |||
 DB 1176 ATCGCTGCTTTCAGTAGAGCTGGGACCCATCCCTGGCTCTCATGTCAGAGATCTTC 1235
 |||
 QY 121 cctctgacatcaagggtgtgctaccgctgctgtctcctcctcagagccctc 180
 |||
 DB 1236 CCTCGACATCAAGGGGTGTGCTACCGGCTGTGTGCTCACCACAGCTGTTTCATGAGCC 1295
 |||
 QY 181 ttctgtgtgaccaaagagtttaacagacatcagagatcctcagagccctc 240
 |||
 DB 1296 TTCTGTGTGACCAAGAGTTTAAACAGCATCATGAGATCCTCAGAGCCCTTACGCGCTTC 1355
 |||
 QY 241 tggctaccgctgctctctgatactcagagctccttcaag 282
 |||
 DB 1356 TGGCTACCGCTGCTCTGTATCTCAGCGCTCTTTTCACG 1397

RESULT 3

LOCUS AB033418 2189 bp mRNA ROD 11-APR-2000
 DEFINITION Rattus norvegicus glut8 mRNA for glucose transporter 8, complete
 cds.
 ACCESSION AB033418
 VERSION AB033418.1 GI:7592743
 KEYWORDS glucose transporter 8.
 SOURCE Rattus norvegicus cDNA to mRNA, clone_11b: testis.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 2189)
 AUTHORS Ishibashi K.
 TITLE Molecular cloning of a new putative glucose transporter
 JOURNAL Published Only in Database (2000) In press
 REFERENCE 2 (bases 1 to 2189)
 AUTHORS Ishibashi K.
 TITLE Direct Submission
 JOURNAL Submitted (08-OCT-1999) to the DDBJ/EMBL/GenBank databases, Kenichi
 Ishibashi, Jichi Medical School, Pharmacology, Minami-Kawachi,
 Kawachi, Tochigi 329-0498, Japan (E-mail: kishibaejichi.ac.jp,
 Tel: 81-285-58-7326, Fax: 81-285-44-5541)

FEATURES

source 1..2189
 /organism="Rattus norvegicus"
 /db_xref="taxon:10118"
 30..1466

gene
/clone_1ib="testis"
109..1548
/gene="glut8"
109..1548
CDS
/gene="glut8"
/note="Glut8 has 12 transmembrane segments"
/product="glucose transporter 8"
/protein_id="BA94383.1"
/db_xref="GI:7592744"
/translation="MSPEDPQETQPLRLSRGAPRGRRVFLATPAALGPLSPFAL
GSSPAIPSLRRTAPALRLGDTAASWFGAVVTLGAAGGIVLGDRLDRGRKLSLL
LCTVPVTFVAVTAAADVWMLLGGRLTGLAGVASLVAPLY1SEIAPVAVKLLS
CYOLAVVTGILLAVAGVLEWRMLAVLGCVPPTLMLLMSYPTPRFLTOHOYE
AMALRFLMSSEGEERPYGAEHQFQALRRPGVHKPLIIGICLIMVPOQLSGVNA
IMEYANTIFEAKFKDSLSAVTGIIVLFTVAVALIMRACRKLILASGIVMFS
MSAFCTFKTOSGPNSSSHVGLIPLVPSAPPAVHGLIALAVGSKLPLIAGVAVGC
PIPLMLSEIFPLHKGATGCVLTNMFMAFLVTKFEFNSIMEILRPYGAFWLTAFC
ILSVLFTLVFPEPKRTLEQITAFHEGR"

BASE COUNT 392 a 702 c 609 g 486 t
ORIGIN

Query Match 100.0%; Score 282; DB 10; Length 2189;
Best Local Similarity 100.0%; Pred. No. 1.2e-59;
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gagcctgctgaltgtcaccctggagcctgctgctgctgctgctgctgctc 60
|||||
Db 1198 GAGCCTCTGTGATGTTACCTGGGCTGGCTGGCTGGCTGAGGACGATGTGCTCTTC 1257
|||||

QY 61 atcgctgcttltgcaagtagctgagcagccatccctgctgctcctcaatgcaagatctc 120
|||||
Db 1258 ATCGCTGGTTTTCAGTAGAGCTGGGACCATGCCCTGGCTCCATGTCAGAGATCTTC 1317
|||||

QY 121 cctctgacatcaagaggtgtgctgctgctgctgctgctcctcaactgctgctgctc 180
|||||
Db 1318 CCTCTGCACATCAAGGCTGTGCTGACCGCTGTGCTGCTCACCACTGCTCATGCTC 1377
|||||

QY 181 ttcttgtagccaaagagtttaacagatcagatgagatccctcaacccctagagccttc 240
|||||
Db 1378 TTCTGTGTACCAAGAGTTTAAACAGATCATGTGAGATCTTCAAGCCTTACGGCCTTC 1437
|||||

QY 241 tgcctcacccgctgctctglatcctcagcgtccctttcacg 282
|||||
Db 1438 TGGCTCACCGCTGCTTCTGTATCCTCAGCGCTCTTTTCACG 1479
|||||

RESULT 4
LOCUS AX076671 2072 bp DNA PAT 06-FEB-2001
DEFINITION Sequence 5 from Patent WO0104145.
ACCESSION AX076671
VERSION AX076671.1 GI:12711202
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2072)
AUTHORS Thorens,B., Ibberson,M. and Uldry,M.
TITLE Glutx polypeptide family and nucleic acids encoding same
JOURNAL Patent: WO 0104145-A 5 18-JAN-2001;
University of Lausanne (CH)
FEATURES
source location/Qualifiers
1..2072
/organism="Mus sp."
/db_xref="taxon:10095"
21..1454
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28497.1"
/db_xref="GI:12711203"

/translation="MSPEDPQETQPLRLPREPARTPRGRVFLASFAALGPLNLFGL
GSSPAIPSLRRTAPALRLGDNAAWFGAVVTLGAAGGIVLGDRLDRGRKLSLL
CTVPVTFVAVTAAADVWMLLGGRLTGLAGVASLVAPLY1SEIAPVAVKLLS
CYOLAVVTGILLAVAGVLEWRMLAVLGCVPPTLMLLMSYPTPRFLTOHOYE
AMALRFLMSSEGEERPYGAEHQFQALRRPGVHKPLIIGICLIMVPOQLSGVNA
IMEYANTIFEAKFKDSLSAVTGIIVLFTVAVALIMRACRKLILASGIVMFS
MSAFCTFKTOSGPNSSSHVGLIPLVPSAPPAVHGLIALAVGSKLPLIAGVAVGC
PIPLMLSEIFPLHKGATGCVLTNMFMAFLVTKFEFNSVEMELRPYGAFWLTAFCAL
SVLFTLVFPEPKRTLEQITAFHEGR"

BASE COUNT 351 a 673 c 584 g 464 t
ORIGIN

Query Match 86.0%; Score 242.6; DB 6; Length 2072;
Best Local Similarity 91.5%; Pred. No. 6.4e-50;
Matches 257; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 gagcctgctgaltgtcaccctggagcctgctgctgctgctgctgctgctc 60
|||||
Db 1104 GAGCCTGTGATGTTACCTGGGCTGGCTGGCTGGCTGAGGACGATGTGCTCTTC 1163
|||||

QY 61 atcgctgcttltgcaagtagctgagcagccatccctgctgctcctcaatgcaagatctc 120
|||||
Db 1164 ATTCCTGGCTTTGCGGTGGCTGGGACCATGCCCTGGCTCTCATGTCAAGATCTTC 1223
|||||

QY 121 cctctgacatcaagaggtgtgctgctgctgctgctgctcctcaactgctgctgctc 180
|||||
Db 1224 CCTCTGCATGTGCAAGGCTGTGCTGACCGCATCTGTGCTCACCACCACTGCTTACGGCC 1283
|||||

QY 181 ttcttgtagccaaagagtttaacagatcagatgagatccctcaacccctagagccttc 240
|||||
Db 1284 TTCTAGTAGACCAAGATTCACACAGCTCATGAGATGTCACACCTTACGGTCTTC 1343
|||||

QY 241 tgcctcacccgctgctctglatcctcagcgtccctttcacg 281
|||||
Db 1344 TGGCTCACCGCTGCTTCTGTGCTGCTCATGCTCCTATTTCAC 1384
|||||

RESULT 5
LOCUS MMU245936 2072 bp mRNA ROD 16-FEB-2000
DEFINITION Mus musculus mRNA for glucose transporter (GLUTX1 gene).
ACCESSION AJ245936
VERSION AJ245936.1 GI:7018606
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2072)
AUTHORS Ibberson,M., Uldry,M. and Thorens,B.
TITLE GLUTX1, a novel mammalian glucose transporter expressed in the
JOURNAL central nervous system and insulin-sensitive tissues
MEDLINE J. Biol. Chem. 275 (7), 4607-4612 (2000)
20138191
REFERENCE 2 (bases 1 to 2072)
AUTHORS Ibberson,M.R.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
CH-1005, SWITZERLAND
FEATURES
source location/Qualifiers
1..2072
/organism="Mus musculus"
/db_xref="taxon:10090"
21..1454
/gene="GLUTX1"
21..1454
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CAB5719.1"
/db_xref="GI:7018607"

BASE COUNT	153 a	323 c	302 g	234 t
ORIGIN	<p> EAADTNGGLAMLAAGNNCLIFACGAVGMPILPMLMSEIFEPPLHKGVAATGCVLTNNF MAPLVTKKEFSLSMEVLRPYGAFWFLASAFICIGVLTLLACVPETNGKLTLEQITTAHFEBOR </p>			
Query Match	78.1%; Score 220.2; DB: 4; Length 1012;			
Best Local Similarity	86.5%; Pred. No. 2,3e-44;			
Matches 243; Conservative	0; Mismatches 38; Indels 0; Gaps 0;			
yy	<p> 1 gaggctgctgagatgctaccctgggagctgagcctgctgctgctgctgctgagcagcatgctcttc 60 </p>			

Chr	Position	Strand	Sequence	Gene
1	10,000,000	+	ATGCGTGGTCTTGCAGTAGAGCTGGAGGACCAACCCCTGCTCCTCATATGCAAGATCTTC	GLUT8
1	10,000,000	-	ATCGCCGGCTTGGCTGCTGGGCTGGGGGGCCCATCCCGTAGCTCTCATATGCAGAGATCTTC	GLUT8
1	10,000,000	+	CCCTGTCACATCAAGAGTGTGTGCTACGGGTGCTGTGCTCTCAACCAACTGTTCAATGACC	GLUT8
1	10,000,000	-	CCTGCTCATGTTCACAGCGGTGGCCACCCGGCTGCTGCTGCATCCACCAACTGGTTCAATGACC	GLUT8
1	10,000,000	+	TTCCTGGTGGACCAAGATTTAAACGACATAGAGATCTCAAGACCTTAAGCGGCTTC	GLUT8
1	10,000,000	-	TTTCTGCTGTGACCAAGATTTCAACGAGTTCATATGAGAGGTCTCAGGCCCTACGGTCTTTC	GLUT8
1	10,000,000	+	TGAGTCAACCGCTGACCTCTGTATCCTCAAGCGTCCCTTTCAAC	GLUT8
1	10,000,000	-	TGGCTGGCCCTCGCCCTCTGCATCTTTCGAGTTCGCGTCTTTTCAAC	GLUT8

ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota: Eutherozoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Metazoa; Primates; Catarrhini; Hominiade; Homo.
TITLE	1 (bases 1 to 1508)
JOURNAL	Doerge,H., Schurmann,A., Bahrenberg,G., Brauers,A. and Joost,H.G. GLUT8, a novel member of the sugar transport facilitator family with glucose transport activity
MEDLINE	J. Biol. Chem. 275 (21), 16275-16280 (2000)
REFERENCE	20283667
AUTHORS	2 (bases 1 to 1508)
TITLE	Joost,H.G.
JOURNAL	Direct Submission
REFERENCE	Submitted (16-JUL-1998) H. Joost, Institute of Pharmacology and Toxicology, Technical University Aachen, Wendlingweg 2, D-52057 Aachen, FRG
FEATURES	Location/Qualifiers
source	1..1508
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/tissue_type="testis"
gene	27..1460
	/gene="GLUT8"
	27..1460
	/gene="GLUT8"
	/codon_start=1
	/product="glucose transporter 8"
	/protein_id="CA889809.1"
	/db_xref="GI:7688146"
CDS	/translation="MPPEPDEPPTPLIGCPGGSAPRRGRVYIFAAAGLPSLGFALLGISPELPISTORAPPAKPRLLDDAASFGAVYITGAAGVLTGMLDRGRKLSLLCSYEPFAGAVITTAQDVWMLLIGRLITGLACGASIVAPYVISEIVPVGRGLSGCVLTVVVGIIILATYLAGVILFMWRVLAVGCVPSIMLLMCEMPTPEFLITQRRDEA
	MAALRFVWGSGEDDEPPIGAEGSFHLALRQPIYKPTIIQVSLAMFOOLSGVNAVM

Db 1208 CTCGCAATGTCACAGGCGGTGCGCAGGACATCTCGTCACCACTGCGCTCATGGCCT 1267
QY 182 ttctgtgtaccacaaagatttaacagcatcgtgagacacacacagccctccttct 241
Db 1268 TTCTCGTGACCAAGAGATTGACGACCTCATGAGGTCCTCAGGCCCTATGAGACCTTCT 1327
QY 242 ggcctacgcgtcctctgtatcctcagcagctccttccac 281
Db 1328 GGCCTGCGCTCCGCTTCTGTCATCTTCAGTGCTTTTCAC 1367

RESULT 12
AX076667 2217 bp DNA PAT 06-FEB-2001
LOCUS Sequence 1 from Patent WO0104145.
ACCESSION AX076667
VERSION AX076667.1 GI:12711198
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2217)
AUTHORS Thorens,B., Ibberson,M. and Oldry,M.
TITLE Glut polypeptide family and nucleic acids encoding same
JOURNAL Patent: WO 0104145-A 1 18-JAN-2001;
University of lausanne (CH)
FEATURES
source Location/Qualifiers
1..2217
/organism="Homo sapiens"
/db_xref="taxon:9606"
348..1781
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28495.1"
/db_xref="GI:12711198"
/translation="MTPEDEETOPLLGPPGCSAPGRVFLAFAAALGPLSEFAL
GSSPAIPSLORAPPAPRLDPAASMFQAVVTGAAAGCVGLGMLVDRGRKLSLL
GSPFVAGFAVITAQDVMMLGRLTLCAGVASLAPVYISLAIAPVRLISLC
VOLMYVYGLLAVIAGVLEHRLAVLGCYPPSLMLLMCFMETPFRLLTOIRROEA
MALRFLWSEGEDEDPPTGADESHLALLRPGITKPTIIGVSLAFQDLQSVNAYM
FAETIIEBAKFRDSSLASVVGVIQVLEFAVALIMDRGRLLLVLSGVVWFS
ACGAYFKLQGGGNSHVAISAPVASQPDASVGLAMLGNMCLFIAGFAVGMCP
PMLMSEIFPLHVKGVATGTCVETNMMLMAFLVLTKEPSSLMELVRLPYGAFWLASAFCTF
SVLPLFCVPEIKGTLEQITAHFEGR"
BASE COUNT 320 a 750 c 672 g 475 t
ORIGIN

Query Match 77.2%; Score 217.6; DB 6; Length 2217;
Best local Similarity 86.1%; Pred. No. 9.5e-44;
Matches 241; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 2 agcctcgtatgttcacatcgtgagcctgctgctgctgctgagcagcagcttctca 61
Db 1432 AGCCTGTGATGCGACGCTGGCGCTGGCTGGCGCTGGCGACATGCTCTTCA 1491
QY 62 tgcctggttttgcaagtagctgaggaacacccctgctccctcagtcagagatctcc 121
Db 1492 TCGCGCGCTTGGCGCTGGCGCGCCATCCCTGCTCCTCATCTCAGAGATCTTCC 1551
QY 122 ctctgcacatcagaggtgtgtctacccgggtctgtgtcctcaccacatgtgtcagcc 181
Db 1552 CTCTGCATGTCAAGCGGTGGCGACGACATCTGCTCTCACCACATGGCTCATGGCCT 1611
QY 182 ttctgtgtacacaaagatttaacagcatcgtgagacacacacagccctccttct 241
Db 1612 TTCTCGGACCAAGAGATTGACGACCTCATGAGGTCCTCAGGCCCTATGAGACCTTCT 1671
QY 242 ggcctacgcgtcctctgtatcctcagcagctccttccac 281
Db 1672 GGCCTGCGCTCCGCTTCTGTCATCTTCAGTGCTTTTCAC 1711

RESULT 13
AL445222
LOCUS AL445222
DEFINITION Human DNA sequence from clone RP11-356B19 on chromosome 9, complete
sequence.
ACCESSION AL445222
VERSION AL445222.9 GI:13810082
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 225370)
AUTHORS Laird,G.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Apr 26, 2001 this sequence version replaced gi:13277497.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; SW:
SWISSPROT; Tr: TrEMBL; Wp: WormPeP; Information on the WormPeP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-356B19 is from the library RP11-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
This sequence is the entire insert of clone RP11-356B19. The true
left end of clone RP11-373J8 is at 96439 in this sequence. The true
right end of clone RP13-225021 is at 9980 in this sequence.
FEATURES
source Location/Qualifiers
1..225370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-356B19"
/clone_1fb="RP11-11.2"
23455..223821
/note="Single clone region. Assembly confirmed by
restriction digest data."
65019..65283
/note="Single clone region. Sequence from reads from a
short insert library derived from a single pUC clone.
Restriction digest data confirm the assembly."
225123..225370
/note="Single clone region. Sequence from reads from a
short insert library derived from a single pUC clone.
Restriction digest data confirm the assembly."
BASE COUNT 56329 a 57092 c 57267 g 54682 t
ORIGIN

Query Match 44.4%; Score 125.2; DB 9; Length 225370;

Query	Subject	Score	Length	Ident	Pos	Gap
QY 52	tgctcttcacatgcgtggtttgcagtagcgtggagaccatccctgctcctcatgtca 111	31.7%	378	6	111	0
Db 61198	TGCTCTCGCTCCACAGCGCTTTGCGGTGGGCTGGGGCCATCCCGCTGCTCATGTCA 61257	62.1%	378	12	111	0
QY 112	gagatcttccctctgcacatacaagggtgtggtctacgggtctgtgtctcaaccactgt 171	0	378	86	171	0
Db 61238	GAGATCTTCCTCTCTCATCTCTCAAGGCGGTGGGAGACAGGACATCTGCTCTCCACCAACTGG 61317	0	378	86	171	0
QY 172	ttcacatgccttctgtgtgcacaaagatttaacagatcatgtagaga tccctcagacctac 231	0	378	86	231	0
Db 61318	CTCATGCGCTTTCTCTGTATACCAAGAGATTTCACAGACGCTCATGTGTAGAGGCGAGGCTCCACC 61377	0	378	86	231	0
QY 232	ggcgcc 237	0	378	86	237	0
Db 61378	AGCACC 61383	0	378	86	237	0
RESULT 14						
LOCUS	AX072418	378 bp	DNA	PAT	25-JAN-2001	
DEFINITION	Sequence 2890 from Patent WO0102568.					
ACCESSION	AX072418					
VERSION	AX072418.1	GI:12582769				
KEYWORDS						
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Enayrola, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumalia; Euthera; Primates; Catarrhini; Homiidae; Homo.					
AUTHORS	Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J., Kassam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D., Lamsan, G., Dmanac, R., Crkajic, R., Dmanac, S., Dickson, M., Labat, I., Leshkowitz, D., Kita, D., Garcia, V. and Strache-Crain, B.					
TITLE	Human genes and gene expression products					
JOURNAL	Patent: WO 0102568-A 2890 11-JAN-2001;					
FEATURES	CHIRON CORPORATION (US); HYSRO, INC. (US)					
Source	Location/Qualifiers					
	1..378					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
BASE COUNT	45 a 130 c 117 g 86 t					
ORIGIN						
Query Match	31.7%; Score 89.4; DB 6; Length 378;					
Best Local Similarity	62.1%; Pred. No. 4, 9e-12;					
Matches 141; Conservative	0; Mismatches 86; Indels 0; Gaps 0					
QY 55	ctcttcacatgcgtggtttgcagtagcgtggagaccatccctgctcctcatgtcaag 114					
Db 31	CTCTTCATCATGCGCTACGCGCTGGGCGTCCCATCCTGCGCTCATGTCTGAG 90					
QY 115	attctccctctgcacatacaagggtgtggtctacgggtctgtgtctcaaccactgt 174					
Db 91	gttcgccccctctgcgccccggtggcggtctacgggtctgtgtctcaaccactgt 150					
QY 175	atgagctctctgtgtgcacaaagatttaacagatcatgtagaga tccctcagacctac 234					
Db 151	AACGCGCTTCTGCTCATACCAAGTCTTCCGCGCAGTGTGTAGACACCTTGGCTCCACAGTG 210					
QY 235	ggctctgcacacgcgtgctctctgtacatcctcagagctccttcaac 281					
Db 211	CTTCTCTCTCTCTCTGCGGCGCATCTGCTGTGTAGAGCTGTGTCTAC 257					
RESULT 15						
LOCUS	HS17803	1580 bp	mRNA	PRI	13-JUL-2001	
DEFINITION	Homo sapiens mRNA for facilitative glucose transporter 6, GLUT6 (SLC2A6 gene).					

	ACCESSION	VJ17803
	VERSION	vj17803.1 GI:9187481
	KEYWORDS	glucose transporter 6, GLUT6; SLC2A6 gene.
	SOURCE	human.
	ORGANISM	Homo sapiens
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
	TITLE	Doege,H., Bocianski,A., Joost,H.G. and Schürmann,A. Activity and genomic organization of human glucose transporter 9 (GLUT9), a novel member of the family of sugar-transport facilitators predominantly expressed in brain and leucocytes Biochem. J. 350 Pt 3, 771-776 (2000)
	JOURNAL	Biochem. J. 350 Pt 3, 771-776 (2000) 20427701
	MEDLINE	2 (bases 1 to 1580)
	REFERENCE	Joost,H.G. Direct Submission Submitted (16-JUL-1998) H. Joost, Institute of Pharmacology and Toxicology , Technical University Aachen, Wendlingweg 2, D-52057 Aachen, FRG
	FEATURES	Location/Qualifiers
	Source	1..1580
	gene	/organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="9" /map="gq34" /cell_type="Leukocytes"
	CDS	47..1570 /gene="SLC2A6" 47..1570 /gene="SLC2A6" /codon_start=1 /product="Facilitative glucose transporter g (GLUT6)" /protein_id="CA896996.1" /db_xref="GI:9187482" /db_xref="SPRMBL:Q9UGQ3" /translation="MOEPLGAEGPDYDTPPEKPPPSBDRARVGTLONRKFATFEFAVEFSGVALVTSPYIPALERSLDPODLHTKSQAISFGSVPTIGAAGLSAMITII NDLGRLSTIFSAPSPAAGCALMAAGHGIMLLGRITFSAGLTRACPYYSEAEAPPVRKALGATPOLMVAVPFSSLTLYALGLELPWMLAVAGAPVLMTILSEPNSRS RELLSRGRDEBALRALAMLRGTDVDVHEEFQIODNNVKROSSVMAREAPHCRPRT TVALLMRLLFOQTGITPILVLIYSIFPSTAVALPKDDAAIVGARVLSTVIALAETMD LAGRKILLEVSAAIMPANLTLGLYIHGGPRPLSNSPAIGLESMWGDAOLAAPAGA YLTIPLLATLMFLFMGVAYCMGPITMIIMSVELPLRARVASGLCVIASMTTAFVLRN SFPLPVSTRGIQVFEPFPFAALCIYSLTFPGCCVFETRGRSLIEGFPMRRSRFLR
	BASE COUNT	216 a 552 c 492 g 320 t
	ORIGIN	
	Query Match	31.7% Score 89.4 DB 9 Length 1580; Best Local Similarity 62.1% Pred. No. 4.3e-12; Matches 141; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
Oy	55 ctctcatcgcgtcgatttcagtagtccggaggaccatcccctgccactccaagtcaagaag 114 	
Db	1256 CTCCTTCATCATNGGCCTACCCTCGTGCGCGGCTGCCATTCCACTCGCTGCTATGTTCTAG 1315	
Oy	115 atctccctctgcacatatcaaagggtgatgacctacggcgatctgltctccacaaccaactgatic 174 	
Db	1316 GTCCCTGCCCGCGCGCGGCGGTGGCTGAGGCCTCACGGGCTCTGCGTCTGGCCAGCTGCGCTC 1375	
Oy	175 atgagcctttctggtgagccaagaagtltaaacagatcatacgaagaaatccccagaccctaagc 234 	
Db	1376 ACCGCTTCGTCTTCACCAACAAGCTTCCTCGCCAGTAGTGAGACCTTCGGCCTCCAGGNG 1435	
Oy	235 gcccttcgagcacacgcgcctcttgtatccctccaacgctccttttac 281 	
Db	1436 CCTTCTCTCTCTTCGCGCGCATCTGTGGTGAGCCGTGCTTTCAC 1482	

Search completed: February 13, 2002, 15:02:46
Job time: 8093 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 13:42:54 ; Search time 283.77 Seconds
(Without alignments)
851.978 Million cell updates/sec

Title: US-09-516-493-11
Perfect score: 282
Sequence: 1 gagcctgctgctgctaccc.....tccctgcgcctccttcacg 282

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues
Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Ceneseq_1101.*
1: /SID2/gcgdata/geneseq/NA1980.DAT.*
2: /SID2/gcgdata/geneseq/NA1981.DAT.*
3: /SID2/gcgdata/geneseq/NA1982.DAT.*
4: /SID2/gcgdata/geneseq/NA1983.DAT.*
5: /SID2/gcgdata/geneseq/NA1984.DAT.*
6: /SID2/gcgdata/geneseq/NA1985.DAT.*
7: /SID2/gcgdata/geneseq/NA1986.DAT.*
8: /SID2/gcgdata/geneseq/NA1987.DAT.*
9: /SID2/gcgdata/geneseq/NA1988.DAT.*
10: /SID2/gcgdata/geneseq/NA1989.DAT.*
11: /SID2/gcgdata/geneseq/NA1990.DAT.*
12: /SID2/gcgdata/geneseq/NA1991.DAT.*
13: /SID2/gcgdata/geneseq/NA1992.DAT.*
14: /SID2/gcgdata/geneseq/NA1993.DAT.*
15: /SID2/gcgdata/geneseq/NA1994.DAT.*
16: /SID2/gcgdata/geneseq/NA1995.DAT.*
17: /SID2/gcgdata/geneseq/NA1996.DAT.*
18: /SID2/gcgdata/geneseq/NA1997.DAT.*
19: /SID2/gcgdata/geneseq/NA1998.DAT.*
20: /SID2/gcgdata/geneseq/NA1999.DAT.*
21: /SID2/gcgdata/geneseq/NA2000.DAT.*
22: /SID2/gcgdata/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	282	100.0	2087	22	AAFS5866 Rat GLUTX1 coding
2	242.6	86.0	2072	22	AAFS5867 Murine GLUTX1 codi
3	219.2	77.7	2080	22	AAPO9552 Human transporter
4	217.6	77.2	2217	22	AAFS5865 Human GLUTX1 codin
5	125.2	44.4	579	21	AA44914 Human secreted exp
6	89.4	31.7	378	22	AAAF67134 Novel human polynu
7	89.4	31.7	823	21	AA23461 CDNA encoding huma
8	89.4	31.7	1577	21	AAAC80569 Human secreted pro
9	89.4	31.7	2011	22	AAFS5871 Rat GLUTX3 coding
10	87.8	31.1	1541	22	AAFS5870 Human GLUTX3 codin
11	74.6	26.5	421	22	AAH9663 Human protein enco

12	64.2	22.8	1675	20	AA232195
13	60.4	21.4	1826	21	AAAC42332
14	54.6	19.4	554	21	AAAC95048
15	53.6	19.0	1395	21	AAAC45857
16	53.6	19.0	1683	21	AAAC39089
17	53.6	19.0	1685	21	AAAC52928
18	53.6	19.0	1741	21	AAAC36319
19	53.6	19.0	1743	21	AAAC52920
20	53.4	18.9	1378	21	AAAC51233
21	53.4	18.9	1642	21	AAAC6009
22	52.6	18.7	1461	22	AAAD12574
23	52.2	18.5	1752	20	AA232203
24	50.8	18.0	718	21	AAAF12490
25	49.4	17.5	2504	22	AAFS5869
26	48	17.0	2471	20	AAFS5516
27	47.2	16.7	2587	12	AAO11148
28	45.4	16.1	1755	21	AAAC0459
29	45	16.0	1815	18	AAAF6495
30	45	16.0	1837	21	AAAC8918
31	45	16.0	2850	20	AA224632
32	45	16.0	2856	21	AAAC5871
33	45	16.0	2893	22	AAFP8714
34	44	15.6	1476	21	AAAF5400
35	43	15.2	2222	22	AAFP3832
36	42.2	15.0	1299	21	AAAF0517
37	41.6	14.8	2592	22	AAAF1396
38	41.2	14.6	1943	18	AAAF6496
39	40.4	14.3	408	22	AAAF67962
40	40.4	14.3	725	22	AAAF67961
41	40.4	14.3	1473	22	AAAF65168
42	40.4	14.3	349980	22	AAAF64966
43	40.2	14.3	1507	21	AAAC3376
44	40.2	14.3	1835	21	AAAC50883
45	40.2	14.3	1837	21	AAAC37492

ALIGNMENTS

RESULT 1	AAFS5866	standard; cDNA: 2087 BP.
ID	AAFS5866	
AC	AAFS5866	
XX		
DT	17-Apr-2001	(first entry)
XX		
DE	Rat GLUTX1 coding sequence.	
XX		
KW	Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;	
KW	hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;	
KW	hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.	
XX		
OS	Rattus sp.	
XX		
PN	WO200104145-A2.	
XX		
PD	18-JAN-2001.	
XX		
PF	14-JUL-2000; 2000WO-IB01042.	
XX		
PR	14-JUL-1999; 99US-0143907.	
PR	27-AUG-1999; 99US-0151140.	
PR	23-FEB-2000; 2000US-0184285.	
PR	13-JUL-2000; 2000US-0616132.	
XX		
PA	(UYLA-) UNIV LAUSANNE.	
XX		
PI	Thorens B, Ibberson M, Uldry M;	
XX	WPI; 2001-112615/12.	
DR	P-PSDB; AAB66933.	
DR		
XX		

Corn hexose carrie
Arabidopsis thalia
Cat flea hindgut a
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Human protein havi
Wheat hexose carl
Aspergillus oryzae
Rat GLUTX2 coding
CDNA encoding fac1
GTP-gene from Chin
Arabidopsis thalia
Arabidopsis thalia
Human glucose tran
Human pancreatic c
Human lung tumor a
Human lung cancer-
Human late stage o
Brevibacterium lac
Human cDNA encodin
Fusarium venenatum
Coding sequence fo
Rat glucose transp
Corynebacterium gl
Corynebacterium gl
C glutamicum codin
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia

PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -

PS Claim 3: Page 71-73; 124pp; English.

XX The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is the coding sequence
 CC for rat GLUTX1.

XX Sequence 2087 BP; 355 A; 673 C; 580 G; 479 T; 0 other;

Query Match 100.0%; Score 282; DB 22; Length 2087;
 Best Local Similarity 100.0%; Pred. No. 2,5e-72;
 Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gagcctgcgtatgttcaacctgggctgctgctgctgctgctgctgctgctgctc 60

DB 1116 gagcctgcgtatgttcaacctgggctgctgctgctgctgctgctgctgctc 1175

QY 61 atgcctggttttcagtagctgagtgagccatccctgctgctcctatgtcaagatcttc 120

DB 1176 atgcctggttttcagtagctgagtgagccatccctgctgctcctatgtcaagatcttc 1235

QY 121 cctctgcacatcaagggtgtgtctacccgctgtgtgtctctcaaccaactggtcattgccc 180

DB 1236 cctctgcacatcaagggtgtgtctacccgctgtgtgtctctcaaccaactggtcattgccc 1295

QY 181 ttctctgtgacccaagaagtttaacagcatatgagatcctcagaccctcagcgcttc 240

DB 1296 ttctctgtgacccaagaagtttaacagcatatgagatcctcagaccctcagcgcttc 1355

QY 241 tggctaccgcgtcctctctatccctcagcgctcctttcacg 282

DB 1356 tggctaccgcgtcctctctatccctcagcgctcctttcacg 1397

RESULT 2

AAFS5867
 ID AAF55867 standard; cDNA: 2072 BP.

XX AAF55867;

DT 17-APR-2001 (first entry)

XX Murine GLUTX1 coding sequence.

KM Murine; GLUTX; gene therapy; vaccine; hexose transport modulator;

KW hexose transport disorder; ischemia; diabetes; hyperglycaemia; ss;

KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.

OS Mus sp.

XX WO200104145-A2.

PD 18-JAN-2001.

PF 14-JUL-2000; 2000MO-IB01042.

PR 14-JUL-1999; 99US-0143907.

PR 27-AUG-1999; 99US-0151140.

PR 23-FEB-2000; 2000US-0184285.

PR 13-JUL-2000; 2000US-0616132.

XX (UYLA-) UNIV LAUSANNE.

PA Thorens B, Ibberson M, Uldry M;

XX WPI: 2001-112615/12.
 DR P-PSDB: AAB66934.

PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -

PS Claim 3: Page 73-74; 124pp; English.

XX The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is the coding sequence
 CC for murine GLUTX1.

XX Sequence 2072 BP; 351 A; 673 C; 584 G; 464 T; 0 other;

Query Match 86.0%; Score 242.6; DB 22; Length 2072;
 Best Local Similarity 91.5%; Pred. No. 6,7e-61;
 Matches 257; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 gagcctgcgtatgttcaacctgggctgctgctgctgctgctgctgctgctgctc 60

DB 1104 gagcctgcgtatgttcaacctgggctgctgctgctgctgctgctgctgctc 1163

QY 61 atgcctggttttcagtagctgagtgagccatccctgctgctcctatgtcaagatcttc 120

DB 1164 atgcctggttttcagtagctgagtgagccatccctgctgctcctatgtcaagatcttc 1223

QY 121 cctctgcacatcaagggtgtgtctacccgctgtgtgtctctcaaccaactggtcattgccc 180

DB 1224 cctctgcacatcaagggtgtgtctacccgctgtgtgtctctcaaccaactggtcattgccc 1283

QY 181 ttctctgtgacccaagaagtttaacagcatatgagatcctcagaccctcagcgcttc 240

DB 1284 ttctctgtgacccaagaagtttaacagcatatgagatcctcagaccctcagcgcttc 1343

QY 241 tggctaccgcgtcctctctatccctcagcgctcctttcacg 281

DB 1344 tggctaccgcgtcctctctatccctcagcgctcctttcacg 1384

RESULT 3

AAD09552
 ID AAD09552 standard; cDNA: 2080 BP.

XX AAD09552;

DT 10-SEP-2001 (first entry)

XX Human transporter and ion channel-1 (TRICH-1) cDNA.

KM Human; transporter and ion channel-1; TRICH-1; vaccine; cystic fibrosis;

KW gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy;

KW hypertension; angina; neurological disorder; asthma; bipolar disorder;

KW dementia; depression; Alzheimer's disease; epilepsy; mood; arrhythmia;

KW Pick's disease; ischemic cerebrovascular disease; AIDS; anxiety; stroke;

KW Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy;

KW demyelinating disease; mental disorder; Schizophrenia; polyomyelitis;

KW muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease;

KW dermatomyositis; diabetes mellitus; immunological disorder; psoriasis;

KW rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus;

KW sickle cell anaemia; Wilson's disease; infertility; Cushing's disease;

KW scleroderma; pulmonary artery stenosis; noctropic; Addison's disease;

KW malabsorption syndrome; hypercholesterolaemia; cancer; ss.

XX Homo sapiens.

OS

FH	Key	Location/Qualifiers
FT	CDS	46..1479
FT		/tag= a
FT		/product= "Human TRICH-1 protein"
XX		
PN	WO200146258-A2.	
XX		
PD	28-JUN-2001.	
XX		
PF	22-DEC-2000; 2000WO-US35095.	
XX		
PR	23-DEC-1999; 98US-0172000.	
PR	14-JAN-2000; 2000US-0176083.	
PR	21-JAN-2000; 2000US-0177332.	
PR	28-JAN-2000; 2000US-0178572.	
PR	02-FEB-2000; 2000US-0179758.	
PR	10-FEB-2000; 2000US-0181625.	
XX		
PA	(INCY-) INCYTE GENOMICS INC.	
XX		
P1	Baughn MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy R, Lal P;	
P1	Hillman JL, Azimzal Y, Yue H, Nguyen DB, Yao MG, Gandhi AR;	
P1	Tang YT, Khan FA.	
XX		
DR	WPI: 2001-418042/4.	
DR	P-PSDB: AAE04888.	
XX		
PT	Novel human transporter and ion channel proteins useful for treating disorders -	
PT	muscle and immunological disorders	
XX		
PS	Claim 5; Page 140-141; 160pp; English.	

CC The present sequence is transporter and ion channel-1 (TRICH-1) cDNA.
CC TRICH is used as vaccine. TRICH is useful for treating a disease or
CC condition associated with decreased expression of functional TRICH,
CC such as transport disorder including amyotrophic lateral sclerosis,
CC cystic fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth
CC disease, Duchenne muscular dystrophy, angina and hypertension,
CC neurological disorders including Alzheimer's disease, amnesia, bipolar
CC disorder, dementia, depression, epilepsy, ischaemic cerebrovascular
CC disease, stroke, cerebral neoplasms, Pick's disease, Huntington's
CC disease and Parkinson's disease, demyelinating diseases, mental disorders
CC including mood, anxiety, schizophrenia and seasonal affective disorder,
CC muscle disorder including cardiomyopathy, myocarditis, polymyositis,
CC dermatomyositis, arrhythmias and asthma and immunological disorders
CC including AIDS, adult respiratory distress syndrome (ARDS), allergies,
CC anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
CC syndrome, systemic lupus erythematosus and other diseases including
CC sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary
CC artery stenosis, Grave's disease, Cushing's disease, Addison's disease,
CC glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers
CC prostatic and viral, bacterial, fungal, helminthic and protozoal
CC infections. TRICH DNA is useful in gene therapy and in diagnostic
CC purposes.
CC
CC
CC Sequence 2080 BP; 300 A; 668 C; 632 G; 460 T; 0 other;
CC
CC

Query Match	77.78%	Score 219.2	DB 22	Length 2080
Best Local Similarity	86.48%	Pred. No. 4.2e-54		
Matches 242: Conservative	0	Mismatches 38	Indels 0	Gaps 0

QY 2 agccgcgcgaggttcaaccgcgggcgcgcgcgtcgtcgtctgaagcaagcaagtcgctcttca 61
Db 1130 agccgccttgatgacgaagcgtgtgggcgcgcgcgcgcgcgcgtgacgcaagtcgctcttca 1189
QY 62 tcgcgcgtgttcgaatgaagctgtgagaccatcccccgcgcgtctcccaatgacaagatctcc 121
Db 1190 tcgcgcgcgtctgagctgtggcgtcggggccatccctcgcgcgtctctctcatgtlagaagatctcc 1249
QY 122 ctctgcacatcaagagtgatgagctacgcggcgtctgtgtctctcaccaactgtlcatagcct 181

Accession	Sequence	Position
Db	1250 cttctgcatgtcaaggcgctggtgagacagcgcatctgcgtctcccaaccaactgtgctatagcct	1309
QY	182 ttctcgtgtgaccaaagaagtttaacagcatcaatggagatctctcagaacctacagcgctctt	241
Db	1310 ttctcgtgtgaccaaagaagtttaacagcatcaatggagatctctcagaacctacagcgctctt	1369
QY	242 ggcctcacgcgtcgtctctgtatcctcagcgtctctttcac	281
Db	1370 ggcctcacgcgtcgtctctgtatcctcagcgtctctttcac	1409

RESULT 4

ID AAF55865 standard; cDNA; 2217 BP.

AC AAF55865;

DT 17-APR-2001 (first entry)

Human GLUTX1 coding sequence.

KW Human; GLUTX; gene therapy; vaccine; hexose transport modulator;

KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.

OS Homo sapiens.

PN W0200104145-A2.

PD 18-JAN-2001.

PF 14-JUL-2000; 2000WO-IB01042.

PR 14-JUL-1999; 99US-0143907.

PR 23-FEB-2000; 2000US-0184285.

XX

XX
XX

XX

DR P-PSDB; AAB66932.

PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in

PT e.g. ischemia and diabetes -

PS Claim 3; Page 70-71; 124pp; English.

CC The present invention relates to GLUTX proteins (AA55865-AA55871 and

CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport

CC treatment of hexose transport disorders such as ischaemia, diabetes,

CC neurodegenerative disease. The present sequence is the coding sequence

XX

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Query Match	77.2%	Score 217.6	DB 22	Length 2217
Best Local Similarity	86.1%	Pred. No. 1.3e-53		
Matches 241	Conservative 0	Mismatches 39	Indels 0	Gaps 0

[illegible]

AAC80569
ID AAC80569 standard; cDNA: 1577 BP.
XX
AC AAC80569;
XX
DT 12-FEB-2001 (first entry)
XX
DE Human secreted protein gene 39 SEQ ID NO:49.
XX
KW Human: secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW neurotropic; neuroprotective; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; autoimmune disease; hyperproliferative disorder;
KW cardiovascular disorder; cerebrovascular disorder; wound healing;
KW nervous system disorder; aging; chemotaxis; ss.
XX
OS Homo sapiens.
XX
PN WO200058467-A1.
XX
PD 05-OCT-2000.
XX
PF 22-MAR-2000; 2000MO-US07505.
XX
PR 26-MAR-1999; 990S-0126502.
XX
PR 17-DEC-1999; 990S-0172410.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
DR WPI; 2000-611712/58.
XX
DR P-PSDB; AAB45158.
XX
PT Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers -
XX
PS Claim 1; Page 371-372; 440pp; English.
XX
CC Polynucleotide sequences AAC80531-C80580 represent cDNA encoding human
CC secreted proteins AAB45120-B45169. Sequences AAB45170-B45225 represent
CC alternative polypeptides encoded by the genes, and amino acid sequences
CC to which they are homologous. The genes and proteins have activities
CC dependent on the tissues and cells in which they are expressed. Examples
CC of their activities include immunosuppressive; antirheumatic;
CC antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
CC cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;
CC fungicide; and ophthalmological. The secreted proteins, polynucleotides,
CC antagonists and agonists may be useful in treating, preventing and/or
CC diagnosing diseases and disorders such as autoimmune diseases
CC e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms
CC of the breast or liver, cardiovascular disorders e.g. cardiac arrest,
CC cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous
CC system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The polypeptides can also be used to aid wound healing and epithelial
CC cell proliferation, to prevent skin aging due to sunburn, to maintain
CC organs before transplantation, for supporting cell culture of primary
CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
CC also be used as a food additive or preservative to increase or decrease
CC storage capabilities. AAC80522-C80530 and AAB45119 represent sequences
CC used in the isolation and characterisation of the genes and proteins of
CC the invention.
XX
XX
SO Sequence 1577 BP; 352 A; 496 C; 445 G; 284 T; 0 other;

Query Match 31.7%; Score 89.4; DB 21; Length 1577;
Best Local Similarity 62.1%; Pred. No. 1.9e-16;
Matches 141; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

OY 55 cttctcatgcgtgttgcagtagctggtggaccatccctcgtgctcctcatgtcagag 114

|||||||
Db 291 cttctcatgcgtgttgcagtagctggtggaccatccctcgtgctcctcatgtcag 350
OY 115 atcttcctctgcacatcaagggtgtggtacacggcgtgtgtccacacactgttc 174
Db 351 gtccctgcctgtgcgtccgtgcgtgcgtgcacagggtcgtgctggtgcacgtgtc 410
OY 175 atggccttctgtgtacacaaagatbaacagatcatgtgagatcctcagacctacggc 234
Db 411 acgcctctgtcctcaccacagctcctcctgtgcaggtgtgtgacacctgcgtccacgg 470
OY 235 gctctcgtgcacgcgtgtcctctgtatcctcaggtccttccac 281
Db 471 ccttctcctctcgcgcacatcgtgtgtgacctgtgttccac 517

RESULT 9
AAF55871
ID AAF55871 standard; cDNA: 2011 BP.
XX
AC AAF55871;
XX
DT 17-APR-2001 (first entry)
XX
DE Rat GLUTX3 coding sequence.
XX
KW Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;
KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS Rattus sp.
XX
PN WO200104145-A2.
XX
PD 18-JAN-2001.
XX
PF 14-JUL-2000; 2000MO-IB01042.
XX
PR 14-JUL-1999; 990S-0143907.
XX
PR 27-AUG-1999; 990S-0151140.
XX
PR 23-FEB-2000; 2000US-0184285.
XX
PR 13-JUL-2000; 2000US-0616132.
XX
PA (UYLA-) UNIV LAUSANNE.
XX
PI Thorens B, Idberson M, Uldry M;
XX
DR WPI; 2001-112615/12.
XX
DR P-PSDB; AAB66938.
XX
PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
PT e.g. ischaemia and diabetes -
XX
PS Claim 3; Page 82-83; 124pp; English.
XX
CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC AAB66932-AAB6941). The GLUTX proteins are related to the facultative
CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischaemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is the coding sequence
CC for rat GLUTX3.
XX
XX
SO Sequence 2011 BP; 337 A; 621 C; 593 G; 459 T; 1 other;

Query Match 31.7%; Score 89.4; DB 22; Length 2011;
Best Local Similarity 62.1%; Pred. No. 2e-16;
Matches 141; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

OY 55 cttctcatgcgtgttgcagtagctggtggaccatccctcgtgctcctcatgtcagag 114

Db 1204 cctctcatatgagcctacgctgagggccacccacccctgctccatcgtcag 1263
 QY 115 atcttcctctcgcacacaaagggtgtgctacggcgctgtgtccaccactgttc 174
 Db 1264 gtctgtccctcgtcgagccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1323
 QY 175 atggccttcctgtgacaaagggttaacagacatcgtgagatcctcagaccctacgc 234
 Db 1324 acagcctctgctcctacacagtaactcctcgctgagtgagtgagccttcaggtg 1383
 QY 235 gccctctgctcagcgctgctctgtatcctcagcgctccttcaac 281
 Db 1384 ccttctctctctcgtcgacatctgctgctcagcctgctctcaac 1430

RESULT 10

AAFS5870
 ID AAF55870 standard; cDNA: 1541 BP.

AC AAF55870;

DT 17-APR-2001 (first entry)

DE Human GLUTX3 coding sequence.

DE Human GLUTX3 gene therapy: vaccine: hexose transport modulator;

KW hexose transport disorder: ischaemia; diabetes; hyperglycaemia; ss;

KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.

OS Homo sapiens.

PN WO200104145-A2.

PD 18-JAN-2001.

PF 14-JUL-2000; 2000WO-1B01042.

PR 14-JUL-1999; 99US-0143907.

PR 27-AUG-1999; 99US-0151140.

PR 23-FEB-2000; 2000US-0184285.

PR 13-JUL-2000; 2000US-0616132.

PA (UUYLA-) UNIV LAUSANNE.

PI Thorens B, Ibberson M, Uldry M;

DR WPT: 2001-112615/12.

DR P-PSDB: AAB66937.

XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in

XX the prevention, diagnosis and treatment of hexose transport disorders,

XX e.g. ischaemia and diabetes -

XX Claim 3; Page 81-82; 124pp; English.

XX The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and

XX AAB66932-AAB66941). The GLUTX proteins are related to the facultative

XX glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport

XX function. The GLUTX proteins may be used in the diagnosis, prevention and

XX treatment of hexose transport disorders such as ischaemia, diabetes,

XX hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a

XX neurodegenerative disease. The present sequence is the coding sequence

XX for human GLUTX3.

Db 1220 cctctcatcagcgctacgcggtgtgctgagggccacccacccctgctccatcgtcag 1279
 QY 115 atcttcctctcgcacacaaagggtgtgctacggcgctgtgtccaccactgttc 174
 Db 1280 gtctgtccctcgtcgagccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1339
 QY 175 atggccttcctgtgacaaagggttaacagacatcgtgagatcctcagaccctacgc 234
 Db 1340 acgctctgctcctcaccacaaagtctcctcgctgagtgagtgagccttcaggtg 1399
 QY 235 gccctctgctcagcgctgctctgtatcctcagcgctccttcaac 281
 Db 1400 ccttctctctctcgtcgacatctgctgctcagcctgctctcaac 1446

RESULT 11

AAH99663
 ID AAH99663 standard; cDNA: 421 BP.

AC AAH99663;

DT 16-OCT-2001 (first entry)

DE Human protein encoding cDNA sequence SEQ ID NO:498.

DE Human: cancer; HIV infection; human immunodeficiency virus;

KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;

KW antitubercial; endocrine; cardiac; central nervous system; vitruide;

KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;

KW antiaggregant; haemostatic; vulnery; antilucer; osteopathic; eczema;

KW dermatological; antiallergic; antisthmatic; antidiabetic; cytostatic;

KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;

KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;

KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;

KW cardiac dysfunction; neuropathology; cardiac anapylaxis; autoimmunity;

KW genetic disease; haematopoietic disorder; platelet disorder; asthma;

KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;

KW allergic rhinitis; diabetes; multiple sclerosis; depression;

KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;

XX neurological disorder; ss.

XX Homo sapiens.

XX WO200153455-A2.

XX 26-JUL-2001.

XX 22-DEC-2000; 2000WO-US35017.

XX 23-DEC-1999; 99US-0471275.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI: 2001-457603/49.

XX P-PSDB: AAM25722.

XX Isolated human polynucleotides encoding polypeptides, useful for the

XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection -

XX Claim 1; Page 564; 1217pp; English.

XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to

XX CC AAM25963. The proteins can have activities based on the tissues and

XX CC cells they are expressed in, such as: antiinflammatory; antirheumatic;

XX CC antiallergic; immunosuppressive; antibacterial; endocrine; cardiac;

XX CC central nervous system; vitruide; anti-HIV; fungicide; antimutagen;

XX CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;

XX CC antilucer; osteopathic; dermatological; antiallergic; antisthmatic;

CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders.
CC
XX
SQ Sequence 421 BP; 72 A; 140 C; 122 G; 87 T; 0 other;

Query Match 26.5%; Score 74.6; DB 22; Length 421;
Best Local Similarity 66.5%; Pred. No. 2.3e-12;
Matches 107; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 34 ctggctctagcagcagctgctctctcatgctggttttgcaaggctgggaccatc 93
DB 1 ctgcgaagccgcaaggctgctctctctgctcagcagctcgtgggtggctccatc 60
QY 94 cctcgctcctcatgctcagagatcttcctctgcacatacaagggtgctgaccgcgtc 153
DB 61 accctggctgctcatgctgagtgctctgcccctgctgctccgctgggctcaggctc 120
QY 154 ttgtctcacaacactggttcagtgctctctctgtagcaca 194
DB 121 tgcgtgctggcagctgctcaccgctctgctcctcaccaa 161

RESULT 12
ID AAZ32195 standard; cDNA; 1675 BP.
XX
AC AAZ32195;
XX
DT 14-JAN-2000 (first entry)
XX
DE Corn hexose carrier protein partial sequence encoding cDNA.
XX
KW Hexose carrier protein; corn; rice; sorghum; soybean; wheat;
KW carbohydrate transport; plant carbon partitioning; manipulation;
XX
XX carbohydrate distribution; ss.
XX
OS Zea mays.
XX
OS
XX
PN WO953082-A2.
XX
PD 21-OCT-1999.
XX
PE 07-APR-1999; 99WO-US07561.
XX
PR 09-APR-1998; 98US-0081131.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Lightner JE, Rafalski JA, Thorpe CJ;
XX
DR WPI: 1999-620438/53.
XX
DR P-PSDB; AAY49625.
XX
PT New hexose carrier proteins used to manipulate carbohydrate transport
XX
PS
XX
CC Claim 2: Page 34-35; 60pp: English.
CC The present sequence encodes part of a hexose carrier protein from the
CC invention, which describes hexose carrier proteins isolated from sorghum,

CC rice, wheat, soybean and corn. Also describes are: (1) a chimeric gene
CC comprising a hexose carrier protein polynucleotide operably linked to
CC regulatory sequences; (2) a transformed host cell comprising the chimeric
CC gene; and (3) altering the level of expression of a hexose carrier
CC protein in a host cell, comprising transforming a host cell with the
CC chimeric gene and growing the cell under expression conditions. Hexose
CC carrier proteins may be used to manipulate carbohydrate transport and to
CC alter whole plant carbon partitioning or to manipulate carbohydrate
XX
XX distribution between cellular compartments.
XX
SQ Sequence 1675 BP; 260 A; 583 C; 503 G; 324 T; 5 other;

Query Match 22.8%; Score 64.2; DB 20; Length 1675;
Best Local Similarity 52.9%; Pred. No. 4e-09;
Matches 138; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 21 ggggctgctgctgctgctgctgtagcagcatgctcccttcacatgctgctgtagcagtagg 80
DB 1319 gggctacgcgcgcgctgctgctgcttcacatgctgctcactgctccgctgctgctgctc 1378
QY 81 ctgggagccatccctgctgctcctcatgccaagagatctccctgcacatacaagggtgt 140
DB 1379 gttgggggcccctggctgctgctgctgctgcgcagagatctcccgctggaatccgcgcgc 1438
QY 141 ggtaccggtctgtgctcctcaccactggttcacatgctgctctctgtgaccaaagatt 200
DB 1439 ggggagagatcaacagctcgtccaacatgcttcacacttcgtacatgcgagagcctt 1498
QY 201 taacagatcatatggaatcctcagaccctcagggcgcctctgctgtaaccgctcttcg 260
DB 1499 cctcaccatgctctgccaactcaagttcgcccttctacttcttcgcggctggtcgt 1558
QY 261 tatcctcagcgtccttcac 281
DB 1559 catcatgaccgctcttatgc 1579

RESULT 13
ID AAC42332
AC AAC42332 standard; DNA; 1826 BP.
XX
AC AAC42332;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35149.
XX
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
OS
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
XX
PR 23-MAR-1999; 99US-0125788.
XX
PR 25-MAR-1999; 99US-0126264.
XX
PR 29-MAR-1999; 99US-0126785.
XX
PR 01-APR-1999; 99US-0127462.
XX
PR 06-APR-1999; 99US-0128234.
XX
PR 08-APR-1999; 99US-0128714.
XX
PR 16-APR-1999; 99US-0129845.
XX
PR 19-APR-1999; 99US-0130077.
XX
PR 21-APR-1999; 99US-0130449.
XX
PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 9905-0130891.
PR 28-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139889.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142970.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.

PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151338.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.

Hybridisation assay; genetic mapping; gene expression control; protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0122180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 23-APR-1999; 990S-0130893.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 06-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0138119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.

PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145091.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147316.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148685.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149375.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 23-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.

PR	10-SEP-1999;	9905-0153070.
PR	13-SEP-1999;	9905-0153758.
PR	15-SEP-1999;	9905-0154018.
PR	16-SEP-1999;	9905-0154039.
PR	20-SEP-1999;	9905-0154779.
PR	22-SEP-1999;	9905-0155139.
PR	23-SEP-1999;	9905-0155486.
PR	24-SEP-1999;	9905-0155659.
PR	28-SEP-1999;	9905-0156458.
PR	29-SEP-1999;	9905-0156596.
PR	04-OCT-1999;	9905-0157117.
PR	05-OCT-1999;	9905-0157753.
PR	06-OCT-1999;	9905-0157865.
PR	07-OCT-1999;	9905-0158029.
PR	08-OCT-1999;	9905-0158232.
PR	12-OCT-1999;	9905-0158369.
PR	13-OCT-1999;	9905-0159293.
PR	13-OCT-1999;	9905-0159294.
PR	13-OCT-1999;	9905-0159295.
PR	14-OCT-1999;	9905-0159329.
PR	14-OCT-1999;	9905-0159330.
PR	14-OCT-1999;	9905-0159331.
PR	14-OCT-1999;	9905-0159637.
PR	14-OCT-1999;	9905-0159638.
PR	18-OCT-1999;	9905-0159584.
PR	21-OCT-1999;	9905-0160741.
PR	21-OCT-1999;	9905-0160767.
PR	21-OCT-1999;	9905-0160768.
PR	21-OCT-1999;	9905-0160770.
PR	21-OCT-1999;	9905-0160814.
PR	21-OCT-1999;	9905-0160815.
PR	22-OCT-1999;	9905-0160880.
PR	22-OCT-1999;	9905-0160981.
PR	22-OCT-1999;	9905-0160989.
PR	25-OCT-1999;	9905-0161404.
PR	25-OCT-1999;	9905-0161405.
PR	25-OCT-1999;	9905-0161406.
PR	26-OCT-1999;	9905-0161359.
PR	26-OCT-1999;	9905-0161360.
PR	26-OCT-1999;	9905-0161361.
PR	28-OCT-1999;	9905-0161920.
PR	28-OCT-1999;	9905-0161992.
PR	28-OCT-1999;	9905-0161993.
PR	29-OCT-1999;	9905-0162142.

Search completed: February 13, 2002, 13:42:56
Job time: 3303 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 13:34:39 ; Search time 92.83 Seconds

(without alignments)
687,997 Million cell updates/sec

Title: US-09-516-493-11

Perfect score: 282

Sequence: 1 gagcgtgtgattcacct.....tccctagcgtcttccacg 282

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*

2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*

3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*

4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*

5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*

6: /cgn2_6/prodata/2/ina/backlitssl.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41.6	14.8	2592	4	US-08-591-025-8
2	36.8	13.0	1365	4	US-08-957-130-12
3	36.8	13.0	2574	2	US-08-677-734A-8
4	36.4	12.9	2818	1	US-08-366-276-1
5	35	12.4	7218	1	US-08-232-463-14
6	33	11.7	2745	1	US-08-363-255-1
7	33	11.7	2745	1	US-08-363-255-13
8	32.6	11.6	1119	4	US-09-178-9738-7
9	32.2	11.4	1635	2	US-08-956-998-1
10	31.6	11.2	2385	1	US-08-393-333-1
11	31.6	11.2	3714	4	US-09-087-465-9
12	31.2	11.1	759	1	US-08-466-603-4
13	31.2	11.1	759	1	US-08-314-503A-4
14	31.2	11.1	759	1	US-08-468-066-4
15	31.2	11.1	759	2	US-08-466-717-4
16	31.2	11.1	759	3	US-08-466-743-4
17	31.2	11.1	759	5	PCT-US95-12414-4
18	31.2	11.1	980	1	US-08-466-603-3
19	31.2	11.1	980	1	US-08-314-503A-3
20	31.2	11.1	980	1	US-08-468-066-3
21	31.2	11.1	980	2	US-08-466-717-3
22	31.2	11.1	980	2	US-08-466-743-3
23	31.2	11.1	980	5	PCT-US95-12414-3
24	31.2	11.1	11288	4	US-08-646-301A-1
25	31.2	11.1	11288	4	US-08-481-968A-4
26	31	11.0	1111	4	US-09-178-9738-9
27	31	11.0	3489	2	US-08-728-323A-1

28	31	11.0	4195	1	US-08-340-011-1	Sequence 1, Appl
29	31	11.0	4195	3	US-08-901-710-1	Sequence 1, Appl
30	31	11.0	4416	3	US-08-795-430-1	Sequence 1, Appl
31	31	11.0	4425	1	US-08-222-616-31	Sequence 31, Appl
32	31	11.0	4425	5	PCT-US95-04228-31	Sequence 31, Appl
33	31	11.0	4795	1	US-08-340-011-3	Sequence 3, Appl
34	31	11.0	4795	3	US-08-901-710-3	Sequence 3, Appl
35	31	11.0	9108	5	PCT-US95-04228-45	Sequence 45, Appl
36	31	11.0	32207	2	US-08-770-379-20	Sequence 20, Appl
37	31	11.0	32207	4	US-08-757-669A-20	Sequence 20, Appl
38	30.6	10.9	2469	3	US-09-087-727-1	Sequence 1, Appl
39	30.4	10.8	524	2	US-08-459-135A-9	Sequence 9, Appl
40	30.4	10.8	524	3	US-08-495-559-9	Sequence 9, Appl
41	30.4	10.8	546	2	US-08-459-135A-5	Sequence 5, Appl
42	30.4	10.8	546	3	US-08-495-559-5	Sequence 5, Appl
43	30.4	10.8	4040	4	US-08-664-962B-1	Sequence 1, Appl
44	30.4	10.8	4040	4	US-09-311-743-1	Sequence 1, Appl
45	30.2	10.7	496	1	US-08-263-413-23	Sequence 23, Appl

ALIGNMENTS

```
RESULT 1
US-09-591-025-8
: Sequence 8, Application US/09591025
: Patent No. 6303373
: GENERAL INFORMATION:
: APPLICANT: Bogan, Jonathan S.
: TITLE OR INVENTION: Method of Measuring Plasma Membrane
: FILE REFERENCE: 0399.1210-004
: CURRENT APPLICATION NUMBER: US/09/591.025
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 60/154,078
: PRIOR FILING DATE: 1999-09-15
: PRIOR APPLICATION NUMBER: 60/138,237
: PRIOR FILING DATE: 1999-06-09
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 8
: LENGTH: 2592
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: modified GLUT4 containing myc tag sequences
US-09-591-025-8
```

Query Match 14.8% Score 41.6: DB 4: Length 2592:
Best Local Similarity 51.0%: Pred. No. 0.011:
Matches 98: Conservative 0: Mismatches 94: Indels 0: Gaps 0:

```
QY 48 catgtgctctcatcgtgtgttgcagtagctggtgacccatccctggtctcat 107
||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 1476 catcttgctgctgtgcatcttttgagatgtgcccctgcccattcttggtcatcgt 1535

QY 108 gtccagatcttccctctgcacatcaagggtgtgtctacgcgggtctgtgcccacaa 167
||| | ||| ||||| | | | | | | | | | | | | | | | | | | |
Db 1536 ggcagagctctcttcagcgggacccgcgcgcagcctgtgtgtgttcttccaa 1595

QY 168 ctgtgtcatgctcttctctgtgaccaaagggttaacagcatatgagatcccaagc 227
||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 1596 ctggaagcacttcatatgtgcagatgtgttccagatgttgcgagagctatggggcc 1655

QY 228 ctacagcgcctt 239
||||| | | |
Db 1656 ctacagcttctct 1667
```

RESULT 2
US-08-957-130-12

RESULT 3
US-08-677-734A-8
Sequence 8, Application US/08677734A
Patent No. 5871919
GENERAL INFORMATION:
APPLICANT: Brant, Steven R.
APPLICANT: Yun, Chris C.H.
APPLICANT: Donowitz, Mark

```

Query Match      13.0%:  Score 36.8;  DB 2;  Length 2574;
Best Local Similarity 51.9%;  Pred.No. 0.22;
Matches 83;  Conservative 0;  Mismatches 77;  Indels 0;  Gaps 0.

QY  116  tcttcctctgacatcaagggtgtggtacggcgctctgtgtcctaccactggttca 175
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db   350  TCTACCTGCTGCCGCCCCATCTGCTGAGACGGCGGTACTTTCATGCCAACCGCTCTTCT 409

QY  176  ttgctctctgtgtgaccaaagaagtttaacagatcatgtagagatccttaacccttagcgcg 235
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db   410  TCGGCAACCTGGGGAGACATCTCTGTGTAGCGCGTCTGTGGGTACCGTGTGGAMACGGGCA 469

QY  236  cctctgtgctaacggtgctctctctgtatccccaagcgtcct 275
      ||  ||  |||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db   470  CCACGGGCGTGTCCCTTACGAGCGCTCTTCTCATGAGGCT 509

RESULT 4
US-08-366-276-1
; Sequence 1, Application US/08366276
; Patent No. 5534409
; GENERAL INFORMATION:
; APPLICANT: Groner, Bernd
; APPLICANT: Gouilleux, Fabrice
; APPLICANT: Makao, Hiroshi
; TITLE OF INVENTION: Cytokine Regulated Transcription Factor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532

```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30B
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/366,276
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9409396.0
: FILING DATE: 11-MAY-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Elmer, James Scoot
: REGISTRATION NUMBER: 36,129
: REFERENCE/DOCKET NUMBER: 4-19992/A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919) 541-8614
: TELEFAX: (919) 541-8689
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2818 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEICAL: NO
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 244..2625
: OTHER INFORMATION: /product= "mammary gland factor"
:
US-08-366-276-1

Query Match          12.9%  Score 36.4; DB 1; Length 2818;
Best Local Similarity 50.6%  Pred. No. 0.29;
Matches 88; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 92 tccctgctccctcatcagagatcttccctctgcacatcaagggtgtgctaacgcgcg 151
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1748 TCCTGTGGCGCGCAGCTGTGGAGGCGCTCAACATGAAATTCAAGCGCGTGCAGAGCA 1807

QY 152 tctgtctccacaaactgtgtcgtcttctgtgtgacaaagagttlaacagcatca 211
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1808 ACCGGGGCCTGCACGAGAGAACCTGTGTTCTGGCGCAGAGCTGTCAACAAACAGCA 1867

QY 212 tggagctctcgaacctaagcgctctcgtgtcaccgctgtcgtatc 265
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1868 GCAGCCACCTCGAGACTACACGCGATGTCTGTCTGTGTCCTCCAGTCAACC 1921

RESULT 5
US-08-232-463-14
: Sequence 14, Application US/08232463
: Patent No. 5670367
: GENERAL INFORMATION:
: APPLICANT: DORNER, F.
: APPLICANT: SCHEIFLINGER, F.
: APPLICANT: FALKNER, F. G.
: TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 1800 diagonal Road, Suite 500
: CITY: Alexandria
: STATE: VA USA
: COUNTRY: USA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/232,463
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/935,313
: FILING DATE:
: APPLICATION NUMBER: EP 91 114 300.6
: FILING DATE: 26-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 30472/114 IMMU
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300
: TELEFAX: (703)683-4109
: TELEX: 899149
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7218 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: CLONE: pTZ9PL-F15
:
US-08-232-463-14

Query Match          12.4%  Score 35; DB 1; Length 7218;
Best Local Similarity 4.5%  Pred. No. 0.85;
Matches 11; Conservative 136; Mismatches 96; Indels 0; Gaps 0;

QY 37 gctgtaggcaacatgtgcctctcctcctcctcgtgttgagtagctgggaaccatccc 96
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1048 GGTGAGGAGACTTGCCGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1107

QY 97 tgcctcctatcagagatcttccctctgcacatcaagggtgtgctaacgcgcgtctgt 156
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1108 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1167

QY 157 gtccctaccacatgctcatgctcttctgtgtagcaacaagagttlaacagcatcagag 216
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1168 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1227

QY 217 atcctcagacctaagcgctctgtgtcctcagcgtcgtctctatcctcagcgtct 276
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1228 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1287

QY 277 ttc 279
    : : :
DB 1288 YYY 1290

RESULT 6
US-08-363-255-1/c
: Sequence 1, Application US/08363255
: Patent No. 5783386
: GENERAL INFORMATION:
: APPLICANT: JACOBS, Jr., WILLIAM R.
: APPLICANT: BLOOM, BARRY R.
: APPLICANT: COLLINS, DESMOND M.
: APPLICANT: de LISLE, GEOFREY W.
: APPLICANT: PASCOPELLA, LISA
: APPLICANT: KAWAKAMI, RIKU P.
: TITLE OF INVENTION: MYCOBACTERIA VIRULENCE FACTORS AND A
: METHOD FOR IDENTIFICATION
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MORRISON & FOERSTER
: STREET: 755 Page Mill Road
: CITY: Palo Alto
: STATE: California
: COUNTRY: USA
```


Query Match	Similarity	11.1%	Score 31.2	DB 1	Length 759
Best Local	Similarity	52.3%	Pred. No.5.6		
Matches	69	Conservative	0	Mismatches	63
				Indels	0
				Gaps	0
Db	150	cgctcgtgtccaccacacactggtatcgtacgtccttcgtgtacacaaagttaacagcat	209		
Db	543	cctcttgcgccctcaccatcgtccgctgtcttcgttttgcgttcgcactcccttcttcacacag	484		
Qy	210	catgagatcctcagaccctcagcgccttcctgtcctcagcgtcgtccttcgtatccctcag	269		
Db	483	ctttcttcttgcgttccttcctcgtatccacttcccccctcattgtaacccttccatctcct	424		
Qy	270	cgctccttcac	281		
Db	423	cctcctcttcac	412		
RESULT	14				
	US-08-468-066-4/c				
	Sequence 4, Application US/08468066				
	Patent No. 5756676				
	GENERAL INFORMATION:				
	APPLICANT: Pasternack, Gary R.				
	APPLICANT: Kuhnada, Francis P.				
	TITLE OF INVENTION: No. 5756676el Mammalian Protein Associated With				
	TITLE OF INVENTION: Uncontrolled Cell Division				
	NUMBER OF SEQUENCES: 9				
	CORRESPONDENCE ADDRESS:				
	ADDRESSEE: Banner, Birch, McKie & Beckett				
	STREET: 1001 G Street, N.W.				
	CITY: Washington, D.C.				
	STATE: District of Columbia				
	COUNTRY: U.S.A.				
	ZIP: 20001				
	COMPUTER READABLE FORM:				
	MEDIUM TYPE: Floppy disk				
	COMPUTER: IBM PC compatible				
	OPERATING SYSTEM: PC-DOS/MS-DOS				
	SOFTWARE: Patent in Release #1.0, Version #1.25				
	CURRENT APPLICATION DATA:				
	APPLICATION NUMBER: US/08/468,066				
	FILING DATE: 06-JUN-1995				
	CLASSIFICATION: 435				
	PRIOR APPLICATION DATA:				
	APPLICATION NUMBER: US 08/314,503				
	FILING DATE:				
	ATTORNEY/AGENT INFORMATION:				
	NAME: Posorske Esq., Laurence H.				
	REGISTRATION NUMBER: 34,698				
	REFERENCE/DOCKET NUMBER: 1107, 47218				
	TELECOMMUNICATION INFORMATION:				
	TELEPHONE: 202 508-9153				
	TELEFAX: 202 508-9299				
	INFORMATION FOR SEQ ID NO: 4:				
	SEQUENCE CHARACTERISTICS:				
	LENGTH: 759 base pairs				
	TYPE: nucleic acid				
	STRANDEDNESS: double				
	TOPOLOGY: linear				
	MOLECULE TYPE: cDNA				
	ORIGINAL SOURCE:				
	ORGANISM: Mus sp				
	FEATURE:				
	NAME/KEY: CDS				
	LOCATION: 3..548				
	US-08-468-066-4				
Query Match	Similarity	11.1%	Score 31.2	DB 1	Length 759
Best Local	Similarity	52.3%	Pred. No.5.6		
Matches	69	Conservative	0	Mismatches	63
				Indels	0
				Gaps	0

```

Oy      150  cgtcgtctccaccacactggtatgatgccttcgttgaccacaagaagtttaacagcat 209
          ||||| ||||| | ||||| | ||||| | ||||| | ||||| | |||||
Db      543  CCTCTTCGCCCTCATCGTCCGGTTCTGTTTTTGCTTGTACTCCCTTCTTTACACAG 484

Oy      210  catgaataacctcaaacctaacygccttctggtcacaccgctctctgtatccatg 269
          ||||| ||||| | ||||| | ||||| | ||||| | ||||| | |||||
Db      483  CTCTCTTTCGTCGCTTCCTGCATCATTCCACTTCCCGCTCATTTGAACCTTCCTACTCTCT 424

Oy      270  cgctccttccaac 281
          | |||| | |||
Db      423  CCTCCTCTCCAC 412

RESULT   15
US-08-466-717-4/C
; Sequence 4, Application US/08466717
; Patent No. 5874234
; GENERAL INFORMATION:
; APPLICANT: Pasternack, Gary R.
; TITLE OF INVENTION: No. 5874234el Mammalian Protein Associated With
; TITLE OF INVENTION: Uncontrolled Cell Division
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: District of Columbia
; COUNTRY: U.S.A.
; ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,717
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/314,503
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Posorske Esq., Laurence H.
REGISTRATION NUMBER: 34,698
REFERENCE/DOCKET NUMBER: 1107,472L8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 508-9153
TELEFAX: 202 508-9299
INFORMATION FOR SEQ ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 759 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
ORIGINAL SOURCE:
ORGANISM: Mus sp
FEATURE:
NAME/KEY: CDS
LOCATION: 3..548
US-08-466-717-4

Query Match 11.1%; Score 31.2; DB 2; Length 759;
Best Local Similarity 52.3%; Pred. No. 5.6;
Matches 69; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
```

Thu Feb 14 07:44:17 2002

us-09-516-493-11.rni

Page 8

0Y 210 celtgagagctccctcagacccctcaggcgccctcttgctcaacgcctgcctctgtaacccag 269
Db 483 cttctctcttcgcctctccctccgctcattccacttcccccgtatttaacccttccgtccct 424
0Y 270 gctcccttttacc 281
Db 423 cctcctcttcacg 412

Search completed: February 13, 2002, 13:34:43
Job time: 2810 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 13:37:40 : Search time 2842.4 Seconds
(without alignments)
1066.109 Million cell updates/sec

Title: US-09-516-493-11
Perfect score: 282
Sequence: 1 ggaacgcgcgatgttcacct.....tccctcagcgtcccttcacg 282

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estpl:*
6: em_estlba:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: gb_est1:*
11: gb_est2:*
12: gb_hlc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rpd:*
20: em_gss_vrl:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	279.4	99.1	319	11	H34451
2	225.6	80.0	823	11	BF120840
3	221.8	78.7	546	10	BE231636
4	219.2	77.7	949	11	BE749509
5	199.4	70.7	934	11	BE910478
6	188.2	66.7	712	11	BE290178
7	171.2	60.7	496	10	AW248655
8	168.8	59.9	440	11	BF774206
9	167.2	59.3	752	10	AL565390
10	154	54.6	538	11	BE672321
11	149.4	53.0	235	11	BF146289
12	149.4	53.0	327	11	BF742266

13	149.2	52.9	917	11	BF140667
14	137.2	48.7	595	11	BG079217
15	108.4	38.4	358	10	AA627408
16	102.8	36.5	295	11	BC303512
17	93.2	33.8	480	10	AW249090
18	89.4	31.7	359	10	AW141319
19	89.4	31.7	684	11	BC325106
20	89.4	31.7	912	11	BF206274
21	88.8	31.5	891	10	AI405787
22	77.4	27.4	911	11	BF311341
23	74	26.2	830	11	BG700749
24	67	23.8	1079	13	AL344131
25	67	23.8	1079	13	CNS0558F
26	66.2	23.5	638	11	BC549214
27	62.2	22.1	289	11	C88387
28	61.8	21.9	279	11	H46443
29	61	21.6	354	11	BI388988
30	60.4	21.4	543	10	BE598634
31	60.4	21.4	555	10	AV524951
32	60.4	21.4	588	10	BE594919
33	60.4	21.4	595	10	BE596754
34	60.4	21.4	654	10	BE594918
35	59.4	21.1	525	10	AW463536
36	59.4	21.1	808	11	BF260051
37	59.2	21.0	855	10	AW940827
38	58.2	20.6	459	10	BE482283
39	56.2	19.9	796	13	AZ985752
40	56.2	19.9	803	10	BE130163
41	56.2	19.9	803	10	BE130163
42	55.8	19.8	487	11	BE878576
43	55.4	19.6	795	13	AQ324535
44	55.2	19.6	521	10	AV522458
45	55.2	19.6	562	10	AV526551

ALIGNMENTS

RESULT 1
LOCUS H34451 319 bp mRNA EST 13-MAR-1998
DEFINITION EST111406 Rat PC-12 cells, NGF-treated (9 days) Rattus sp. cDNA 5' end similar to Glucose transporter, mRNA sequence.

ACCESSION H34451 GI:979868
VERSION H34451
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 319)

AUTHORS Lee,N.H., Weinstock,K.G., Kirkness,E.F., Earle-Hughes,J.A., Fuldner,R.A., Marmaras,S., Glodek,A., Gocayne,J.D., Adams,M.D., Kerlavage,A.R., Fraser,C.M. and Venter,J.C.
Comparative expressed-sequence-tag analysis of differential gene expression profiles in PC-12 cells before and after nerve growth factor treatment

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 92, 8303-8307 (1995)

MEDLINE 95396786

COMMENT Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
For clone availability please contact the TIGR Database (tigrinfo@db.tigr.org)
Seq primer: M13 Reverse.

FEATURES Location/Qualifiers
1..319

source /organism="Rattus sp." /db_xref="ATCC (inost):2005510"

/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
BASE COUNT 76 a 182 c 162 g 124 t 2 others
ORIGIN

Query Match 78.7%; Score 221.8; DB 10; Length 546;
Best Local Similarity 86.8%; Pred. No. 2,3e-46;
Matches 244; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

OY 1 gagctgtgatagttcacctgggctgctgctgtgctgtaagcagcatgtcctcttc 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2 GACCCGCCGCTCCAGCGGCGGCTGGCTGGCTGGCGGCGGCGGCTGGCTGGCTTC 61
OY 61 atgcgtgttttcagtaggctggggaccatccctggcttcctcagtaagaattctc 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 62 ATTGCTGGCTTCGCCCTGGGCTGGGCGCATCCCTGGCTCCCATATGTCAGATCTTC 121
OY 121 cctctcacatcaagggtgtgctacccgctgtgtctcaccacatgttcatagcc 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 122 CCTCTGCACCTCAAGGCGGCGGCGGCGGCTGTCTCTCTCACTGCTCATGGCC 181
OY 181 ttctgttgaccagaagtttaacagcatcatgagatcctcagacctacgctcttc 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 182 TTCTCTGTGACCAAGGATTCAAGCAGCTCATGAGGTGCTCAGGCCCTACGGTCTTC 241
OY 241 ttgtcaccgctgctctgtatcctcagctcttcttcac 281
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 242 TGCTCGCTCCGCTTCTGATCTTCAGTCTCTTTTAC 282

RESULT 4
LOCUS BG749509 949 bp mRNA EST 15-MAY-2001
DEFINITION 6020707648F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:484441 5',
mRNA sequence.
ACCESSION BG749509
VERSION BG749509.1 GI:14060162
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 949)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LCM1681 row: p column: 12
High quality sequence stop: 745.

FEATURES

Source 1..949
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:484441"
/clone_lib="NIH_MGC_43"
/issue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: This is a NIH-MGC Library. |"
BASE COUNT 160 a 302 c 264 g 223 t
ORIGIN

Query Match 77.7%; Score 219.2; DB 11; Length 949;
Best Local Similarity 86.4%; Pred. No. 1.2e-45;
Matches 242; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

OY 2 agctgtctgatagttcacctgggctgctgctgtgctgtaagcagcatgtcctctca 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 81 AGCCTGTGATGCCAGCGCTGGGCGGCTGGCTGGCGGCGGCGGCGGCTGGCTTC 140
OY 62 tgcgtgttttcagtaggctggggaccatccctggcttcctcagtaagaattctc 121
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 141 TCGCCGCGCTTGGCGGCTGGGCGGCGGCGGCGGCGGCTGGCTTCATGTCAGATCTTC 200
OY 122 cctctcacatcaagggtgtgctacccgctgtgtctcaccacatgttcatagcct 181
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 201 CTCTGCATGTCAGGCGGCGGCGGCGGCGGCGGCTGTCTCTCACTGCTCATGGCC 260
OY 182 ttctgttgaccagaagtttaacagcatcatgagatcctcagacctacgctcttc 241
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 261 TTCTCTGTGACCAAGGATTCAAGCAGCTCATGAGGTGCTCAGGCCCTACGGTCTTC 320
OY 242 ggtcaccgctgctctgtatcctcagctcttcttcac 281
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 321 GCGTTCCTCCGCTTCTGATCTTCAGTCTCTTTTAC 360

RESULT 5
LOCUS BE910478 934 bp mRNA EST 20-OCT-2000
DEFINITION 601501005F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3903048 5',
mRNA sequence.
ACCESSION BE910478
VERSION BE910478.1 GI:10407108
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 934)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9706 row: i column: 01
High quality sequence stop: 710.

FEATURES

Source 1..934
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3903048"
/clone_lib="NIH_MGC_70"
/issue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT
Average insert size 1.1 kb. Library constructed by Life
Technologies." |"
BASE COUNT 143 a 291 c 287 g 213 t
ORIGIN

RESULT	10
BG672321	
LOCUS	BG672321
DEFINITION	BG672321 538 bp mRNA
ACCESSION	D98CDD10 Rat DRG library Rattus norvegicus EST
VERSION	BG672321.1 mrna sequence.
FEATURES	CDNA clone D98CDD10 5'
KEYWORDS	BG672321.1 GI:13894420

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
<i>Rattus norvegicus</i>	1 (bases 1 to 538)	Xiao,H.S., Han,Z.G., Zhang,F.X., Huang,Q.H., Lu,Y.J., Bao,L., Fu,G., Guo,C., Yan,Q., Jin,S.X., Zhu,Z.D., Xu,X.R., Li,N.G., Chen,Z. and Zhang,X.	Distinct gene expression profiles of rat dorsal root ganglion induced by peripheral nerve axotomy	Unpublished (2001)	Contact: Zhang Xu

TITLE	JOURNAL	COMMENT
Distinct gene expression profiles of rat dorsal root ganglion induced by peripheral nerve axotomy unpublished (2001)		
Contact: Zhang Xu Laboratory of Sensory System Institute of Neuroscience 320 Yue Yang Road, Shanghai 200031, P.R.China Tel: 86-21-64748700-121 Fax: 86-21-64713446 Email: xu.zhang@ion.ac.cn		
This clone is also available at Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhanjiang Hi-Tech Park (xu.zhang@ion.ac.cn) or Han Zeguang (hanzg@chgc.sh.cn)		
PCR Primers FORWARD: T3 BACKWARD: T7 Seq Primer: T3 POLYA=No.		

```

FEATURES
source
    Location/Qualifiers
1.538
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="DRNCPD10"
/clone_1ib="Rat DRG Library"
/sex="male"
/rissue_type="dorsal root ganglion"
/dev_stge="adult"
BASE COUNT
118 a 146 c 138 g 136 t
ORIGIN

```

Query Match	54.68;	Score 154;	DB 11;	Length 538
Best Local Similarity	85.18;	Pred. No. 3.6e-29;		

	Matches	172:	Conservative	0:	Mismatches	30:	Indels	0:	Gaps	0:
QY	81	ctggaggaccatccctctgctcctcaatgtaagatcttcctctctgacataaggagt	140							
Db	1	cttgggagaccattccctggctcttattttcaaaacatttcctttgcacattaaaggggg	60							
QY	141	ggtcccgagctctgtgtccctaccacatggtatcatgcttcttggtgaccaaagatt	200							
Db	61	gggtaccggggtctggggctccacacactggttcatgaggctttctggagaccmaaagatt	120							
QY	201	taacagatcatgtagatcctcagaacctagagcagcctctcttgctaccagctgctctg	260							
Db	121	taacagcatcatatgagaaatcttaaacctctacggcgcccttttgacctcaagggtcttttg	180							
QY	261	tatccatcagctcctctttacag	282							
Db	181	tattcttagcgcccttttcacg	202							

RESULT	11
LOCUS	BF146289
DEFINITION	BF146289 235 bp mRNA
ACCESSION	EST00181 rabbit blastocyst mRNA to cDNA
VERSION	clone 81 similar to glucose transporter 8 (GLUT8), mRNA sequence
KEYWORDS	BF146289.1 GI:11027684
SOURCE	EST.
	rabbit.

ORGANISM	Euryctolagus cuniculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus. 1 (bases 1 to 235)
REFERENCE AUTHORS TITLE	Kietz S., Augustin R. and Fischer B. Expression of glucose transporters in rabbit preimplantation embryos Unpublished (2001)
JOURNAL COMMENT	Contact: Kietz S Department of Anatomy and Cell Biology Martin Luther University Halle-Wittenberg, Medical Faculty Grosse Steinstrasse 52, D-06097 Halle, GERMANY Email: sliike.kietz@gmx.de Seq primer: 7' Forward High quality sequence stop: 235 POLYA=No.

FEATURES	Location/Qualifiers
source	1..235
	/organism="Oryctolagus cuniculus"
	/strain="hybrid strain zika"
	/db_xref="taxon:9986"
	/clone="c81"
	/clone_1lb="rabdit blastocyst mRNA to cDNA"
	/dev_stage="six days old preimplantation embryo"
BASE COUNT	39 a 78 c 61 g 57 t
ORIGIN	

	Query Match	Best Local Similarity	53.0%;	Score 149.4;	DB 11;	Length 235;
	Matches 162;	Conservative	0;	Mismatches 21;	Indels 0;	Gaps 0;
QY	99	gclctccatbltcagagatctccctccctgcacataaggtgtgctccggcgtctgtt	158			
Db	1	gcttctcatgtcagaaatcttcccgctgcacctcCAAGGgtgtgctcAAGGgctcttgcgcg	60			
QY	159	ccctccacaactcgttcaatgcgcttcttgltgaccnaaagatlltaacagcatatgagat	218			
Db	61	ccctcaccacacttgctctcatggcttccTCTAGTACCAGAGgttCAGcgcctCATtGAGGc	120			
QY	219	ccctcagacctaagcgagccttctgtctcaacgcctgcgctcttgtatacctaagcgtcttt	278			
Db	121	ctctccggcccttACGGGgcttcttgcttgccttgccttgccttTGCATtTCAGGccttctt	180			
QY	279	cac 281				

RESULT 14
 BG079217 595 bp mRNA
 LOCUS H3037C03-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
 DEFINITION H3037C03 5', mRNA sequence.
 ACCESSION BG079217
 VERSION BG079217.1 GI:12561785
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 595)
 AUTHORS Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
 T.S., Carter,M.G. and Ko,M.S.H.
 TITLE Verification and initial annotation of NIA mouse 15K cDNA clone set
 JOURNAL Unpublished (2001)
 COMMENT other_ESTS: H3037C03-3
 CONTACT: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@gsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please
 visit <http://19sun.grc.nia.nih.gov/cDNA/15K.html> for details.
 Plate: H3037 row: C column: 03
 Seq primer: -21M13 Reverse
 High quality sequence stop: 595
 POLYA-No.

FEATURES

SOURCE

Location/Qualifiers
 1..595
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="H3037C03"
 /clone_lib="NIA Mouse 15K cDNA Clone Set"
 /sex="Clones arrayed from a variety of cDNA libraries"
 /dev_stage="Clones arrayed from a variety of cDNA libraries"
 /lab_host="DH10B"
 /note="Vector: pSPORT1; Site1: SalI; Site2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT) expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

BASE COUNT 95 a 177 c 182 g 141 t
 ORIGIN

Query Match 48.7%; Score 137.2; DB 11; Length 595;
 Best Local Similarity 91.8%; Pred. No. 6,6e-25;
 Matches 145; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 gagcgtcatatcttaccctggagcgtgctgctgtagcagcagctgctc 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 438 GAGCCTGTGATGTCCAGATGGAGACTGGCTGTGCTGATGAGCAGATGTCTCTTC 497
 QY 61 atgcgtgcttctgagtagcgtggagccatccctgctcctcctcagagcttc 120
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 498 ATTGCTGGCTTGGCGGTGGGAGCCCAACCCCTGCTCATGTGAGATCTTC 557
 QY 121 cctctgacataaagggtgtggtaccggcgtgctgt 158
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 558 CCTCTCATGTCAAGGAGGTGTGCTACCGGACATCTGTGT 595

RESULT 15
 AA627408 358 bp mRNA
 LOCUS AA627408/c
 DEFINITION ng496d04.sl CGAP Co10 Homo sapiens cDNA clone IMAGE:1147207 3' similar to TR:G1209756 G1209756 INTEGRAL MEMBRANE PROTEIN.;, mRNA sequence.
 ACCESSION AA627408.1 GI:2539503
 VERSION AA627408
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 358)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmett-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bdrp/image/image.html
 Insert Length: 507 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 349.

FEATURES

SOURCE

Location/Qualifiers
 1..358
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1147207"
 /clone_lib="NCI-CGAP-Co10"
 /tissue_type="Colon tumor RER+"
 /lab_host="DH10B"
 /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from RER+ colon tumor, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo (N-Soares4)."

BASE COUNT 87 a 104 c 101 g 66 t
 ORIGIN

Query Match 38.4%; Score 108.4; DB 10; Length 358;
 Best Local Similarity 82.7%; Pred. No. 1.2e-17;
 Matches 124; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 84 ggagccatccctgctcctcatatgacagatcttcctctgcacatcaagagtgtgc 143
 ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 355 GGGGCCATCCCTCTGCTCTCATGTCAAGATCTTCCTGTGATGTCAAGGGGTGC 296
 QY 144 taccggcgtctgttctcaccacaactgtctcagccttctggtgaccacaagattaa 203
 ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 295 GACAGCAGATCTCGTCACCAACACACACACACACACACACACACACACACAC 236
 QY 204 cagcatcatgagagatccctcagacacctcagg 233
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 235 GAGCTCATGTCTCGTCTGTGAG 206

Thu Feb 14 07:44:19 2002

us-09-516-493-11.rst

Page 9

Search completed: February 13, 2002, 13:37:42
Job time: 2989 sec

Thu Feb 14 07:45:08 2002

us-09-516-493-9.rst


```

FT      Domain                               /label= Transmembrane_domain
FT      438..457
XX      /label= Transmembrane_domain
XX      MO200146258-A2.
XX      28-JUN-2001.
XX      22-DEC-2000; 2000WO-US35095.
XX      23-DEC-1999; 99US-0172000.
XX      14-JAN-2000; 2000US-0176083.
XX      21-JAN-2000; 2000US-0177332.
XX      28-JAN-2000; 2000US-0178572.
XX      02-FEB-2000; 2000US-0179758.
XX      10-FEB-2000; 2000US-0181625.
XX      (INCY-) INCYTE GENOMICS INC.
XX      Baughn MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy R, Lal P;
XX      Hillman JL, Azimzal Y, Yue H, Nguyen DB, Yao MG, Gandhi AR;
XX      Tang YF, Khan FA;
XX      WPI; 2001-418042/44.
XX      N-PSDB; AAD09552.
XX      Novel human transporter and ion channel proteins useful for treating
XX      and preventing transport, neurological, muscle and immunological
XX      disorders
XX      Claim 1; Page 112-113; 160pp; English.
XX
XX      The present sequence is transporter and ion channel-1 (TRICH-1) protein.
XX      TRICH is used as vaccine. TRICH is useful for treating a disease or
XX      condition associated with decreased expression of functional TRICH,
XX      such as transport disorder including amyotrophic lateral sclerosis,
XX      cystic fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth
XX      disease, Duchenne muscular dystrophy, angina and hypertension,
XX      neurological disorders including Alzheimer's disease, amnesia, bipolar
XX      disorder, dementia, depression, epilepsy, ischemic cerebrovascular
XX      disease, stroke, cerebral neoplasms, Pick's disease, Huntington's
XX      disease and Parkinson's disease, demyelinating diseases, mental disorders
XX      including mood, anxiety, schizophrenia and seasonal affective disorder,
XX      muscle disorder including cardiomyopathy, myocarditis, polymyositis,
XX      dermatomyositis, arhythmias and asthma and immunological disorders,
XX      including AIDS, adult respiratory distress syndrome (ARDS), allergies,
XX      anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
XX      syndrome, systemic lupus erythematosus and other diseases including
XX      sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary
XX      artery stenosis, grave's disease, Cushing's disease, Addison's disease,
XX      glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers
XX      psoriasis and viral, bacterial, fungal, helminthic and protozoal
XX      infections. TRICH DNA is useful in gene therapy and in diagnostic
XX      purposes.
XX      Sequence 477 AA:
XX
XX      Query Match 99.4%; Score 2310.5; DB 22; Length 477;
XX      Best Local Similarity 99.6%; Pred No. 4.3e-224;
XX      Matches 452; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
XX
XX      1 RRVFLAARFALGPIISFGFALGYSSPAIPSLORAPPAARLDDAASWFGAVVTGAAG 60
XX      24 rrvflaaraafalgpiisfgfalgyspsaipslorappaarlddaaswfgavvtlgaag 83
XX      61 GVLGMLVDRAGRKSLILLCSPFYAGRAVITAADVMMLGRLTLGACVASTAVAPV 120
XX      84 gvlgmldvdragrksllilcspfyagravitaadvmmlgrrltlgacvastavapv 143
XX      121 YISETIAPVAVNRGLSCVOLMNVGILLAYLAWLLEWRMLAVIGCCVPPSLMLLMCMFP 180
XX      144 yisetiavpavnrglscvolmnavgillaylawnlewrmlavigccvppslmlilmcmfp 203

```

```

QY      181 EPRFELTQHRROEAL-ALRFLMGSSEQWEDPPTAEQSFHLALLROPCTIKPPIIGVSL 239
XX      |||||
XX      204 etprflltqhrrgeamaarlltwsgsgwedppigaegsfhlallrpglykpfllgysl 263
QY      240 MAFQQLSGVMNMFYAEITFEFAKFKDSSLASVWGVGVQVLEFVVAALIMRARRLIIV 299
XX      |||||
XX      264 mafqqlsgvmnmfyaeitfeefakfkdsllasvvgvlgvltftavaalimdragrlliv 323
QY      300 LSGVMVFSTSAFGAFYFLTLTGCGPQNSHVAISAPVSAQPDVASGLAMTAVGSMCLFIA 359
XX      |||||
XX      324 lsgvmvfstsaafgayfklitg9pqnshvasapvaeagpvaevglawlavsgmclfia 383
QY      360 GFVAGWGPPIWMLMSEIFPLHVKGVATGICVLTWMLMAFLVTKEFSSLMELVLRPYCAFVL 419
XX      |||||
XX      384 gfavgwgpplwmlmseifplhvkgvatgicvltwmlmaflvtkefsslmelvlpayafvl 443
QY      420 ASARCFISVLTFLFCVPEPTKGTLEQITAHFEGR 453
XX      |||||
XX      444 asarcfislvtflfcvpetkgtleqitahfeqr 477

```

```

RESULT 2
AAB66932
ID      AAB66932 standard; Protein; 477 AA.
XX
XX      AAB66932:
XX
XX      17-APR-2001 (first entry)
XX
XX      Human GLUTX1.
XX
XX      Human: GLUTX; gene therapy; vaccine; hexose transport modulator;
XX      hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
XX      hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX      Homo sapiens.
XX      MO200104145-A2.
XX
XX      18-JAN-2001.
XX
XX      14-JUL-2000; 2000WO-IB01042.
XX
XX      14-JUL-1999; 99US-0143907.
XX      27-AUG-1999; 99US-0151140.
XX      23-FEB-2000; 2000US-0184285.
XX      13-JUL-2000; 2000US-0616132.
XX
XX      (UWLA-) UNIV LAUSANNE.
XX
XX      Thorens B, Ibberson M, Uldry M;
XX
XX      WPI; 2001-112615/12.
XX      N-PSDB; AAF55865.
XX
XX      Nucleic acids encoding GLUTX glucose transporter proteins, useful in
XX      the prevention, diagnosis and treatment of hexose transport disorders,
XX      e.g. ischaemia and diabetes -
XX
XX      Claim 11; Page 70-71; 124pp; English.
XX
XX      The present invention relates to GLUTX proteins (AAF55865-AAF55871) and
XX      AAB66932-AAB6941). The GLUTX proteins are related to the facilitative
XX      glucose carriers GLUT1-GLUT5 and have hexose binding and/or facilitative
XX      function. The GLUTX proteins may be used in the diagnosis, prevention and
XX      treatment of hexose transport disorders such as ischaemia, diabetes,
XX      hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
XX      neurodegenerative disease. The present sequence is human GLUTX1.
XX
XX      Sequence 477 AA:

```

Query Match 99.0%; Score 2301.5; DB 22; Length 477;
 Best Local Similarity 99.1%; Pred. No. 3.4e-223;
 Matches 450; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

1 RRFLLAFAAALGPIISFGFALGYSPPAIPSLORAPAPRLDDAAASWFGAVVTLGAAG 60
 |||||
 Db 24 rrvflaafaaalgpisfgfalgysspaipslqraapaprlldaaaswfgavvltgaag 83
 61 GVLGWLVDNRAGRKLSLLCSVPFVAGFAVITTAODVWMLLGRLLTGLACGVAASLAPV 120
 |||||
 Db 84 gvlgwlvdragrklsllcsvpfvagfavitaaodvwmllgrlltglacgyaslvapy 143
 121 YISEIAPVAVRGLLGSCVQIMVVGILLAYLAGVLEWRMLAVLCVPPSLMLLKCMP 180
 |||||
 Db 144 yiseiapvavrgllgscvqlmvvgillaylagvlewrlavlgcvpslmllmcimp 203
 181 ETPRFLLTHORROEAI-ALRFLLMGSEOGWEDDPIGAEOSFHLLALRQPGIYKPFITGVSL 239
 |||||
 Db 204 etprfllthqrqeaamaalrlwgseegweddpiigaegsfhlaillrpgiylkprllgvsll 263
 240 MAFOQLSGVNAVVFYAEITFEERAFKDDSLASVVGVIQVLTFAVAALLMDRAGRLLLV 259
 |||||
 Db 264 mofqqlsgvnavvmfyaeitfeekfkddslasvvgviqvlftavaallmdragrlllv 323
 300 LSGVWVWFSTSAFGATFKLTGGPGNSHVAISAPVSAOPVDASVGLAWLAVGSMCLFT 359
 |||||
 Db 324 lsgvwmvfstsaftgayfklitggpgnshvaaisapvsadqvdasyglawlavgsmclft 383
 360 GFAVWGPIPLMLMSEIFPLHVKGVATGICVLTNMLMAFLVTKEFSSLMVLRPYGAFWL 419
 |||||
 Db 384 gfavwgpiplmlmseifplhkvatgicvltmwlmalfvtkefsslmvltprygaftw 443
 420 ASAFICFVSVLFTLCVPEETKGTLEQITAHFEGR 453
 |||||
 Db 444 asaficfsvlftlcvcpeetkgtleqitahfeqr 477

RESULT 3
 AAB66939
 ID AAB66939 standard; Protein: 478 AA.
 XX
 AC AAB66939;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE GLUTX1 consensus sequence.
 XX
 KW GLUTX; gene therapy; vaccine; hexose transport modulator; human; rat;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia; murine;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX
 OS Homo sapiens.
 OS Mus sp.
 OS Rattus sp.
 XX
 PN WO200104145-A2.
 XX
 PD 18-JAN-2001.
 XX
 PF 14-JUL-2000; 2000MO-1B01042.
 XX
 PR 14-JUL-1999; 9905-0143907.
 PR 27-AUG-1999; 9905-0151140.
 PR 23-FEB-2000; 2000US-0184285.
 PR 13-JUL-2000; 2000US-0616132.
 XX
 PA (UYLA-) UNIV LAUSANNE.
 XX
 PI Thorens B, Ibberson M, Uldry M;
 XX
 DR WPL; 2001-112615/12.
 XX
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in

PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischaemia and diabetes -
 XX
 PS Claim 11; Page 74-75; 124pp; English.
 XX
 CC The present invention relates to GLUTX proteins (AA55865-AA55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is a consensus sequence
 CC for GLUTX1.
 XX
 SQ Sequence 478 AA;

Query Match 90.5%; Score 2104; DB 22; Length 478;
 Best Local Similarity 89.0%; Pred. No. 2.7e-203;
 Matches 405; Conservative 22; Mismatches 26; Indels 2; Gaps 2;

1 RRFLLAFAAALGPIISFGFALGYSPPAIPSLORAPAPRLDDAAASWFGAVVTLGAAG 60
 |||||
 Db 24 rrvflaafaaalgpisfgfalgysspaipslrtcapallrgdaaswfgavvltgaag 83
 61 GVLGWLVDNRAGRKLSLLCSVPFVAGFAVITTAODVWMLLGRLLTGLACGVAASLAPV 120
 |||||
 Db 84 gvlgwlvdragrklsllcsvpfvagfavitaaodvwmllgrlltglacgyaslvapy 143
 121 YISEIAPVAVRGLLGSCVQIMVVGILLAYLAGVLEWRMLAVLCVPPSLMLLKCMP 180
 |||||
 Db 144 yiseiapvavrgllgscvqlmvvgillaylagvlewrlavlgcvpslmllmcimp 203
 181 ETPRFLLTHORROEAI-ALRFLLMGSEOGWEDDPIGAE-OSFHLLALRQPGIYKPFITGVSL 238
 |||||
 Db 204 etprfllthqrqeaamaalrlwgseegweddpiigaegsfhlaillrpgiylkprllgvsll 263
 239 LMAFOQLSGVNAVVFYAEITFEERAFKDDSLASVVGVIQVLTFAVAALLMDRAGRLLLV 298
 |||||
 Db 264 lmfqqlsgvnavvmfyaeitfeekfkddslasvvgviqvlftavaallmdragrlllv 323
 299 VLSGVWVWFSTSAFGATFKLTGGPGNSHVAISAPVSAOPVDASVGLAWLAVGSMCLFT 358
 |||||
 Db 324 vlsgvwmvfstsaftgayfklitggpgnshvaaisapvsadqvdasyglawlavgsmclft 383
 359 GFAVWGPIPLMLMSEIFPLHVKGVATGICVLTNMLMAFLVTKEFSSLMVLRPYGAFWL 418
 |||||
 Db 384 gfavwgpiplmlmseifplhkvatgicvltmwlmalfvtkefsslmvltprygaftw 443
 419 LSAFICFVSVLFTLCVPEETKGTLEQITAHFEGR 453
 |||||
 Db 444 lsaaficfsvlftlcvcpeetkgtleqitahfeqr 478

RESULT 4
 AAB66939
 ID AAB66939 standard; Protein: 478 AA.
 XX
 AC AAB66939;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE GLUTX1.
 XX
 KW Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX
 OS Rattus sp.
 OS
 XX
 PN WO200104145-A2.
 XX
 PD 18-JAN-2001.

XX 14-JUL-2000; 2000MO-IB01042.
PF
XX
PR 14-JUL-1999; 990US-0143907.
PR 27-AUG-1999; 990US-0151140.
PR 23-FEB-2000; 2000US-0184285.
PR 13-JUL-2000; 2000US-0616132.
XX
XX (UYLA-) UNITV LAUSANNE.
PA
PI Thorens B, Ibberson M, Uldry M;
XX
XX WPI; 2001-112615/12.
DR N-PSDB; AAF53866.
XX
XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
PT e.g. ischemia and diabetes -
PT
XX
XX Claim 11; Page 71-72; 124pp; English.
XX
XX
XX The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is rat GLUTX1.
XX
XX Sequence 478 AA;

Query Match	88.5%	Score 2057	DB 22	Length 478
Best Local Similarity	85.7%	Pred. No. 1.4e+198		
Matches 390	Conservative 32	Mismatches 31	Indels 2	Gaps 2

OY	1	RRVFIAAFAAALGPIRSEFEALGYSSPILPSIORAPAPAPRIIDDAASAFCAVWTLGAAG	60
Dd	24	rvfiactaaaiiprstrgfafaygspasprtrrappaltlgtddaasvfgaavvllgaaq	83
OY	61	GVLGWMLVDRAGRKLSLLELCSVPVAFVITTAADVMMLGKRLTLGLACGVASLVA	120
Dd	84	gvlgwmlldragrkrlsllelcvprlvgtfavitaaardwmljgfrlltqlagyaaslva	143
OY	121	YISELAAYAVAGCIGLSCQQLMNVVGGILLAYLAGVLEWRLALVIGCPSSLMLLMK	180
Dd	144	yiseiaaypavavgllygscqqlnmvvggillayagavlewrtaavlgvcpssllmlm	203
OY	181	ETPRFLTLQHRROEAI-ALRFLMCSDEQWEDPRIGAE-OSFHALLRQSPITK	238
Dd	204	etprflltlqhqygemaalrflwgsseegweeprygaethgfgqlamrlrpyvhp	263
OY	239	LMAEQULSGVAWVWFYAEETIEEKKFKDSSLASVVGVOYLTAVALIMDRAGR	298
Dd	264	lmvfqqlsgvaalmfyaatlfeekfkfddsslasvvgvlyltafavaalimdragr	323
OY	299	VLSGVWVFSTASAGAVYKRLQGGPGNSNHAIAIPASAPVDAVSGLAWLAWGSK	358
Dd	324	alsgyimvfmsaafgtyrklqgspnsnshvgllvplsaepadchgladavaysmc	383
OY	359	AGFAVGMPIWMLMSEIFPLHVKGVATGICVLTNMLMAELVTKFSSSIMEVL	418
Dd	384	agfavgmprpwllmseifphikgvatgvcltnwfmafivtkcfssimeilrpyafw	443
OY	419	LASAFCLFSVLFTEPCVPETKGRKLEQITAFEBE	453
Dd	444	ltaaclisvylflltlypetkyrqlqgltatfegr	478

RESULT	5
AAB66934	
ID	AAB66934 standard; Protein; 477 AA.
XX	

AC	AAB66934;
XX	
DT	17-APR-2001 (first entry)
XX	
DE	Murine GLUTX1.
XX	
KW	Murine; GLUTX; gene therapy; vaccine; hexose transport modulator; hexose transport disorder; ischaemia; diabetes; hyperglycaemia; hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
KV	
OS	Mus sp.
XX	
PM	WO200104145-A2.
XX	
PD	18-JAN-2001.
XX	
PF	14-JUL-2000; 2000WO-1B01042.
XX	
PR	14-JUL-1999; 99US-0143907. 27-AUG-1999; 99US-0151140. 23-FEB-2000; 2000US-0184285. 13-JUL-2000; 2000US-0616132.
PR	
XX	(UYLA-) UNIV LAUSANNE.
PA	
XX	
PI	Thorens B, Ibberson M, Uldry M;
XX	
DR	WPI: 2001-112615/12. N-PDB: AAF55867.
XX	
PT	Nucleic acids encoding GLUTX glucose transporter proteins, useful in the prevention, diagnosis and treatment of hexose transport disorders, e.g. ischemia and diabetes -
PT	
XX	
PS	Claim 11; Page 73-74; 124pp: English.
XX	
CC	The present invention relates to GLUTX proteins (AAF55865-AAF55871 and AAB66932-AAB66941). The GLUTX proteins are related to the facultative glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport function. The GLUTX proteins may be used in the diagnosis, prevention and treatment of hexose transport disorders such as ischaemia, diabetes, hyperglycemia, hypoglycaemia, a glucose metabolism disorder and/or a neurodegenerative disease. The present sequence is murine GLUTX1.
CC	
XX	
SQ	Sequence 477 AA;
Query Match	87.8%; Score 2041.5; DB 22; Length 477;
Best Local Similarity	85.9%, Pred. No. 5,3e-197;
Matches 391; Conservative 32; Mismatches 29; Indels 3; Gaps 3	

Query Match	Similarity	87.88	Score	20.41.5	DB	22	Length	477
Best Local	Similarity	85.9%	Pred.	0.53e-197				
Matches	Conservative	32	Mismatches	29	Indels	3	Gaps	
QY	1	RRVFLAAFAAALGPIFLSFGALGYSPPALPSLQORAAPAPRLDDAASWFGAVTTLGAAG	60					
Db	24	rrvflaaafaalgpfnlfgalgysspalpsrlrtlapellrldgnaaswfgavtvlgaag	83					
QY	61	GYLGGWIMDRAGRKSLILLCGVPFVAGRAVITTAADVMMLLGGHLLGLIAGVASTAPV	120					
Db	84	gllgwllldlsgfkrslillevprflvgavltaardwmlilggrillglgaqvastlapy	143					
QY	121	YISLSEYPAVAVGILGSCVQLVWVVGILAVLAGVLEWRMLAVGCVPSLMLLMCEMP	180					
Db	144	yiselaipavvglllygscqqlmvvgillayagavlewrmlavlgcvpslmlllmcymp	209					
QY	161	ETPRFLTLQHRROEARALRFLMKSEGCWEDMPIGAEOSFHALLRKQGYIKFFIIGVS	238					
Db	204	etprfltlcqhyqgeamaalrlfwseegeewpvgaeqhgqlalrlrrpjlkykplligis	263					
QY	239	LMAFQOLSGVNAVMPYATETFEFKFKROSSIASVVGCTOVLPAFAVALIMDRAGRLL	298					
Db	264	lmvlgqlsgvnailmlyansltfeekfkrossiasvvtvgllqylflavaaalmdragrlll	322					
QY	299	VLSGVAVVFSTSAFCAYFKLLQGGPGNSSHVALPAPASQCPVDASVGLIAMLAVSGMLFT	358					

Db 324 alysvimfmsafytlflltssnshvyl-vpiaepvdyvqqlawlavgsmclli 382
 QY 359 ACFAVGWGPIMWLMSEIFPLHVKVATGICVLTNWLMALFLTKRFSIMEVLRPYGAFW 418
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 383 agfavgwgpipplinselfplhvkvatgicvltwfmalfvtkefnsveml:pygalw 442
 QY - 419 IASAFICFSVLTLCVPEPTKGTLEQITAHFEGR 453
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 443 ltaafcalsvltfltvpektgkrltqevlahfeqr 477
 RESULT 6
 AAE06579
 ID AAE06579 standard; Protein: 262 AA.
 AC AAE06579;
 XX
 DT 25-SEP-2001 (first entry)
 DE Human protein having hydrophobic domain, HP10784.
 XX
 KW Human; hydrophobic domain; gene therapy; nutritional supplement;
 KW cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;
 KW multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;
 KW haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;
 KW Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;
 KW haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
 KW contraceptive; antifertility; antinflammatory.
 KW
 OS Homo sapiens.
 XX
 PN WO200149728-A2.
 PD 12-JUL-2001.
 XX
 DE 28-DEC-2000; 2000WO-JP093359.
 PF
 PR 06-JAN-2000; 2000JP-0000585.
 PR 06-JAN-2000; 2000JP-0000588.
 PR 11-JAN-2000; 2000JP-0002895.
 PR 03-FEB-2000; 2000JP-0026862.
 PR 03-MAR-2000; 2000JP-0058367.
 PR
 XX (PROT-) PROTEGENE INC.
 PA (SAGA) SAGAMI CHEM RES CENT.
 XX
 PI Kato S, Kimura T;
 DR WPI: 2001-418355/44.
 DR N-PSDB: AAD12574.
 XX
 PT Human proteins with hydrophobic domains and the nucleic acids encoding
 PT them, useful for preventing diagnosing and treating e.g. cancer,
 PT Alzheimer's and inflammation -
 PT Alzheimer's and inflammation -
 PS Claim 1: Page 75; 563pp; English.
 XX
 CC The present sequence is human protein with hydrophobic domain,
 CC HP10784. The polynucleotide and polypeptide of the invention
 CC may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate polypeptide expression. The polynucleotides
 CC may be used to produce the polypeptide, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. The
 CC polynucleotides and its complementary sequences may also be used as DNA
 CC probes in diagnostic assays and also used in gene therapy. The
 CC polypeptides may also be used as antigens in the production of antibodies
 CC and in assays to identify modulators of polypeptide expression and
 CC activity. The polypeptides and nucleic acids may be used as nutritional
 CC supplements, to modulate cytokine and cell proliferation activity, to
 CC modulate immune stimulation or suppression (e.g. for the treatment of
 CC microbial infections and autoimmune disorders such as multiple sclerosis,
 CC rheumatoid arthritis and insulin-dependent diabetes), to modulate
 CC haematopoiesis, to modulate tissue growth activity (e.g. for the

CC treatment of Parkinson's disease, Huntington's disease and Alzheimer's
 CC disease), to modulate activin and inhibin activity (e.g. for controlling
 CC fertility), to modulate chemotactic and chemokinetic activity, to
 CC modulate haemostatic and thrombolytic activity, to modulate receptor
 CC ligand activity, to modulate inflammation and to inhibit tumour growth.
 XX
 SQ Sequence 262 AA:
 Query Match 43.4%; Score 1007.5; DB 22; Length 262;
 Best Local Similarity 90.5%; Pred No. 3.2e-93;
 Matches 200; Conservative 2; Mismatches 6; Indels 13; Gaps 1;
 QY 1 RRVFLAFAAALGPLSPGALGYSSPAIPSLQRAAPPAPRLDDAASWEGAVYTIGAAC 60
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 24 rrvflaafaalqplstfgalgysspaipslgraappaprlddaaswfgavvltigaag 83
 QY 61 GVLGWLVDPRAGKRLSLLCVSFPVAGFAVITTAQDVWMLLGGRLTLGLACVAAVLAPY 120
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 84 gylgwlvdpragrklslllcsvpfvagfavltlaagdvmllygrrlltqlacgyasvlavp 143
 QY 121 YTSEIAYPAVRGLGSCVOLMVVGTLLAYLACWLEWRLAVLGCVPSSLMLLMCFMP 180
 CC ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 144 ylseiaypavrgllgscvqlmvvvglllaylaagwlewtvawlgcvpslmllmcmfp 203
 QY 181 ETFRFLTQHRRQEAR-----ALRFLMGSEQGW 208
 CC ||||||||||||||||
 Db 204 etpfrlltqhrrgeaapylvrghyqheclrrlllgadpgw 244
 RESULT 7
 AAB66938
 ID AAB66938 standard; Protein: 503 AA.
 XX
 AC AAB66938;
 XX
 DT 17-APR-2001 (first entry)
 DE
 DE Rat GLUTX3.
 DE
 XX Rat; GLUTX: gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 KW
 OS Rattus sp.
 XX
 PN WO200104145-A2.
 PD 18-JAN-2001.
 XX
 DE 14-JUL-2000; 2000WO-IB01042.
 PF
 PR 14-JUL-1999; 99US-0143907.
 PR 27-AUG-1999; 99US-0151140.
 PR 23-FEB-2000; 2000US-0184285.
 PR 13-JUL-2000; 2000US-0616132.
 PR
 XX (UYLA-) UNIV LAUSANNE.
 PA
 PI Thorens B, Ibberson M, Uldry M;
 DR WPI: 2001-112615/12.
 DR N-PSDB: AAF55871.
 XX
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischaemia and diabetes -
 PT e.g. ischaemia and diabetes -
 PS Claim 11: Page 82-83; 124pp; English.
 XX
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport

CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischemia, diabetes,
CC hyperlipaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is rat GLUTX3.

SQ Sequence 503 AA;

Query Match	41.0%	Score 953	DB 22	Length 503
Best Local Similarity	44.28%	Pred. No. 2.4e 87		
Matches 205	Conservative 79	Mismatches 168	Indels 12	Gaps 6

Oy 1 RRVFLAAFAAALGLSPLECEALGYSSPAPISIORAAPAPRILDDAAABAFGVNVLGAAG 60
 Db 35 rrvflataaafagnfsglgaivlvcsyipaikrissdalltdkigaafsgsvflgaag 94
 Oy 61 GVLGWLVDRAVGRKLSLLECGVPFVAGFAVITAAADVMMLGRLNLTGLACGVASTLAPV 120
 Db 95 glsamlindllgrkrisfmsafpaigaylmagarglwmlllgmtlgfsgltaacip 154
 Oy 121 YISEIAYPAVNGLLGSCVOLMWWVGLITAYIAGVLEMRILAVIGCVPPSLMLLMCFMP 180
 Db 155 yiseiaypavngvgaigatpcqllnavfsgsistylglllpwrwlaavegspvilmllisfmp 214
 Oy 181 ETPRLEIQRHROEAIARFLM---GSGOGEDPPIC---AEOSFHALLR--OPGITYK 232
 Db 215 nsprlllskrssdeaelgailwlradesvhefegidnvrqgsyrsweaeeweyrtp 273
 Oy 233 FFIGVSLMAFOOLSGVNAVMFYAETIFEEAK-FKDDSLASVYNGVIOFLPFAVALIMDR 291
 Db 274 illelvmlfqlgltgltipillyltdlstdstsvlpsqgdaaigavrlsvllaavlmld 333
 Oy 292 AGRLLIALLGSCVWVWFNSARGAFAFKLTCG--GCGNSSHAIVASIPVSAQPVDSVGLAML 349
 Db 334 agrkwllyvssimfvaalltglylvyqdlvprtlpnstveivltlggtccqpaafnylcll 393
 Oy 350 AVGSMCLFPIAGFAVWGCDIPWILMSEIFPLHVKGCATGICVLTNMLMAFLTFEFSLME 409
 Db 394 plalmclfmjngamvgpwtclmsevplrarqvasgclvswlatafalktyllavn 453
 Oy 410 VLRPGAFWLASARCFISVLTFLTECEVBEYTKGKLTLEQTAAFEGR 453
 Db 454 afqlqvprfffsaalcslslttgccevyetrsgtseqleaafhntr 497

RESULT	8
AAB66941	
ID	AAB66941 standard; Protein; 507 AA.

AC	AAB66941;
XX	
DT	17-APR-2001 (first entry)
XX	
DE	GLUTX3 consensus sequence.
XX	
KW	GLUTX: gene therapy; vaccine; hexose transport modulator; human; rat;
KW	hexose transport disorder; ischemia; diabetes; hyperglycemia;
KX	hypoglycemia; glucose metabolism disorder; neurodegenerative disease
XX	
OS	Homo sapiens.
OS	Rattus sp.
XX	
PN	WO200104145-A2.
XX	
PD	18-JAN-2001.
XX	
PF	14-JUL-2000; 2000WO-1B01042.
XX	
PR	14-JUL-1999; 99US-0143907.
PR	27-AUG-1999; 99US-0151140.
PR	23-FEB-2000; 2000US-0184285.
PR	13-JUL-2000; 2000US-0616132.
XX	

PA (UYLA-) UNIV LAUSANNE.
XX
XX Thorens B, Ibberson M, Uldry M,
PI
XX WPI; 2001-112615/12.
DR

PT Nucleic acids encoding GLUTx glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
PT e.g. ischemia and diabetes -
XX
PS Claim 11: Page 83-84; 124pp; English.

CC The present invention relates to GLUTX proteins (AAE55865-AAE55871 and
CC AAB66332-AAB66941). The GLUTX proteins are related to the facultative
CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischaemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is a consensus sequence
CC for GLUTX3.

Sequence 507 AA;

Query Match	40.0%;	Score 930;	DB 22;	Length 507;
Best Local Similarly	44.2%;	Pred. No. 5e-85;		
Matches 209; Conservative	78;	Mismatches 160;	Indels 26;	Gaps 10

[illegible]

RESULT	9
AAB66937	
ID	AAB66937 standard; Protein; 507 AA

AC	AAB66937;	
XX		
DT	17-Apr-2001	(first entry)
XX		
DE	Human GLUTX3.	
XX		
KW	Human; GLUTX; gene therapy; vaccine; hexose transport modulator	
TX	hexose transport disorder; ischemia; diabetes; hyperglycaemia;	

PR	18-AUG-1999;	990S-0149426
PR	20-AUG-1999;	990S-0149722
PR	20-AUG-1999;	990S-0149723
PR	20-AUG-1999;	990S-0149920
PR	23-AUG-1999;	990S-0149929
PR	23-AUG-1999;	990S-0149930
PR	25-AUG-1999;	990S-0150566
PR	26-AUG-1999;	990S-0150884
PR	27-AUG-1999;	990S-0151065
PR	27-AUG-1999;	990S-0151066
PR	27-AUG-1999;	990S-0151080
PR	30-AUG-1999;	990S-0151303
PR	31-AUG-1999;	990S-0151438
PR	01-SEP-1999;	990S-0151930
PR	07-SEP-1999;	990S-0152363
PR	10-SEP-1999;	990S-0153070
PR	13-SEP-1999;	990S-0153758
PR	15-SEP-1999;	990S-0154018
PR	16-SEP-1999;	990S-0154039
PR	20-SEP-1999;	990S-0154779
PR	22-SEP-1999;	990S-0155139
PR	23-SEP-1999;	990S-0155486
PR	24-SEP-1999;	990S-0155659
PR	28-SEP-1999;	990S-0156458
PR	29-SEP-1999;	990S-0156596
PR	04-OCT-1999;	990S-0157117
PR	05-OCT-1999;	990S-0157753
PR	06-OCT-1999;	990S-0157865
PR	07-OCT-1999;	990S-0158029
PR	08-OCT-1999;	990S-0158232
PR	12-OCT-1999;	990S-0158369
PR	13-OCT-1999;	990S-0159293
PR	13-OCT-1999;	990S-0159294
PR	13-OCT-1999;	990S-0159295
PR	14-OCT-1999;	990S-0159329
PR	14-OCT-1999;	990S-0159330
PR	14-OCT-1999;	990S-0159331
PR	14-OCT-1999;	990S-0159637
PR	14-OCT-1999;	990S-0159638
PR	18-OCT-1999;	990S-0159584
PR	21-OCT-1999;	990S-0160741
PR	21-OCT-1999;	990S-0160767
PR	21-OCT-1999;	990S-0160768
PR	21-OCT-1999;	990S-0160770
PR	21-OCT-1999;	990S-0160814
PR	21-OCT-1999;	990S-0160815
PR	22-OCT-1999;	990S-0160980
PR	22-OCT-1999;	990S-0160981
PR	22-OCT-1999;	990S-0160989
PR	25-OCT-1999;	990S-0161404
PR	25-OCT-1999;	990S-0161405
PR	25-OCT-1999;	990S-0161406
PR	26-OCT-1999;	990S-0161359
PR	26-OCT-1999;	990S-0161360
PR	26-OCT-1999;	990S-0161361
PR	28-OCT-1999;	990S-0161920
PR	28-OCT-1999;	990S-0161992
PR	28-OCT-1999;	990S-0161993
PR	29-OCT-1999;	990S-0162142

	Query Match	28.3%	Score 657;	DB 21;	Length 437;
	Best Local Similarity	36.1%;	Pred. No. 1,3e-57;		
	Matches 166; Conservative	82;	Mismatches 170;	Indels	42; Gaps 10;
QY	3 VFLAAFAALAPLSPGFGALGYSSPAIPSLORAPPARPRLDDAASWPFACVATYLGAAGCV	62			
Dd	2 YLSLTFVCGPFPATGSCAGSYSPA---GAIRNDLSITIAEFSILYSILTIYGMAIGAI	57			
QY	63 LGGMLVDRAGRKLLSLLCSVPFVAGFAVITTAODPWMLTGGRIILLTGACGASLVAPRYI	122			
Dd	58 TSGPIADIVYRKGMARYSSATCVCVGWIALIFAKGVVALDIDRGLATCYMGAFISYVVPFI	117			
QY	123 SEIAYPAVRGLISSCVOLMNVVGIILAYLAGWLEWRMLAVLGCVPSIMLLMCMPBET	182			

Db 118 aeiapktfrgaltlnqilicgsvsfligltlwrvtaliglipcaasflgffipes 177
QY 183 PRFLLTQHNRQE-ATLARFLMGSEGWEDPPIGAEQSFLLALLR-----QPGIV 230
Db 178 prwlakvgdrfdeafatfrlrgkkadisee--aaeiqdyietlerlpkakmldlqrryi 235
QY 231 KPFIIGVSLMAEQQLSGVAAVMFYAETTFEEAKFKDSSIASVVGVYIQLFTFAVALIND 290
Db 236 rsvllaftgmivqfvgigigiclyssiteqgfp-rlrigmiiyavliqvitalnaplvd 294
QY 291 RAGRLLLVLS--GVYMWSTSAFCAYFKLTGGGPGNSSHVALISAPVSAOPVDASVGLAW 348
Db 295 ragrfpdlivsatglivgcllaavsflylkv-----hdmaheav-----pv----- 334
QY 349 LAVGSMCIETAGFAVAGWGFIPWLLMSEIFPLMHYKGVATGICVLTNMLAFIVTKPESSIM 408
Db 335 lavvgimviygsfsgmgmpwvmselrplnlkgyagmatlvnwfgaavsvytrflm 394
QY 409 EVLRPYGAFWLASAFCIFSVLFTLPCVPETKCKTLEQITFA 448
Db 395 s-wssygtlliyaalnaiaivlaivpetkqkltleqiga 433

RESULT 11
AAG13749
ID AAG13749 standard; Protein: 447 AA.
XX AAG13749;
AC AAG13749;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 13358.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 13358.
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132407.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135253.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.

[illegible]

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

```

PR 13-OCT-1999; 99US-0159293-
PR 13-OCT-1999; 99US-0159294-
PR 13-OCT-1999; 99US-0159295-
PR 14-OCT-1999; 99US-0159329-
PR 14-OCT-1999; 99US-0159330-
PR 14-OCT-1999; 99US-0159331-
PR 14-OCT-1999; 99US-0159337-
PR 14-OCT-1999; 99US-0159637-
PR 18-OCT-1999; 99US-0159584-
PR 18-OCT-1999; 99US-0160741-
PR 21-OCT-1999; 99US-0160767-
PR 21-OCT-1999; 99US-0160768-
PR 21-OCT-1999; 99US-0160770-
PR 21-OCT-1999; 99US-0160814-
PR 22-OCT-1999; 99US-0160815-
PR 22-OCT-1999; 99US-0160980-
PR 22-OCT-1999; 99US-0160981-
PR 25-OCT-1999; 99US-0160989-
PR 25-OCT-1999; 99US-0161404-
PR 25-OCT-1999; 99US-0161405-
PR 26-OCT-1999; 99US-0161359-
PR 26-OCT-1999; 99US-0161360-
PR 26-OCT-1999; 99US-0161361-
PR 28-OCT-1999; 99US-0161920-
PR 28-OCT-1999; 99US-0161921-
PR 28-OCT-1999; 99US-0161992-
PR 29-OCT-1999; 99US-0162142-

```

Query Match 28.3%; Score 657; DB 21; Length 463;
 Best Local Similarity 36.1%; Pred. No. 1,4e-57;

Matches 166; Conservative 82; Mismatches 170; Indels 42; Gaps 10;

```

QY 3 VELAAFAALGPLSFALGYSSPAIPSLQRAAPPAPRLDDAASWFGAVYTLGMAAGV 62
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 28 YLSTFVAVCPGFafgscagyspa---gaaindisltlaetslfgslttfgamgai 83
QY 63 LSGMLVDAGRKSLILCSYFVFAVITTAQOVMMILGRLITGLACGVASLVAPEYI 122
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 84 tsgpiadlvgtkgamrvssafcvvgwlaillfakgvvaldigrilatgyymgaftsyvvpflfi 143
QY 123 SEIAYPAVRGILGSCVDIMVVGILAYLAGVLEWRMLAVLCVPSLMLIMCEMPET 182
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 144 aelapktfrgalttlnqlllctgysvfflgtlvtwvialigilipcaastlglfifpes 203
QY 183 PRELLTQHRROE-ALALRFLMGSEQGWEDDPPIGAESGFHALLR-----QPgiY 230
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 204 ptwlaakvgrdtfealtrkirkkadisee--aaeiqdyietlerlpkxkmlatffrryl 261
QY 231 KPFLIGVSLMAFOOLSGVNAMVFAETTFPEAKFKDSSLASVVGVIQVLTAVAILMD 290
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 262 rsvlatfglmvfqfgfnglclftslstteqagf-ptrlgmillyavilqavvltainpdivd 320
QY 291 RAGRILLVLVS--GVNVVESTSAFGAYFKLTOGGPGNSHVAISAPVSAQPDVASGLAW 348
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 321 ragrkplllvsatgvlvgcllaavsfylkv-----hdmahav-----pv----- 360
QY 349 LAVGSMCLFTAGFAVNGCPPLMLMSIRPLHLKGAATGTCVLTNNMLMALVTKKESSLM 408
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 361 lavvgimvlygstsaqgamvwmselfpinlkyyagagmatlwnfgawavaytlnflm 420
QY 409 EVLRPYGAFMLASAFICFVSLFTLFCVPEETKGTLEQITA 448
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421 s-wssygtfllyaainalavfialvtpckkyktlseqiga 459

```

RESULT 13

ID AAG20978 standard; Protein; 437 AA.

AC AAG20978;

XX 17-OCT-2000 (first entry)

```

XX XX Arabidopsis thaliana protein fragment SEQ ID NO: 23367.
DE Protein identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX EP1033405-A2.
XX PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825-
XX 05-MAR-1999; 99US-0123180-
XX 09-MAR-1999; 99US-0123548-
XX 23-MAR-1999; 99US-0125788-
XX 25-MAR-1999; 99US-0126284-
XX 29-MAR-1999; 99US-0126785-
XX 01-APR-1999; 99US-0127462-
XX 06-APR-1999; 99US-0128234-
XX 08-APR-1999; 99US-0128714-
XX 16-APR-1999; 99US-0129845-
XX 19-APR-1999; 99US-0130077-
XX 21-APR-1999; 99US-0130449-
XX 23-APR-1999; 99US-0130510-
XX 23-APR-1999; 99US-0130891-
XX 28-APR-1999; 99US-0131449-
XX 30-APR-1999; 99US-0132048-
XX 30-APR-1999; 99US-0132407-
XX 04-MAY-1999; 99US-0132484-
XX 05-MAY-1999; 99US-0132485-
XX 06-MAY-1999; 99US-0132486-
XX 06-MAY-1999; 99US-0132487-
XX 07-MAY-1999; 99US-0132863-
XX 11-MAY-1999; 99US-0134256-
XX 14-MAY-1999; 99US-0134218-
XX 14-MAY-1999; 99US-0134219-
XX 14-MAY-1999; 99US-0134221-
XX 14-MAY-1999; 99US-0134370-
XX 18-MAY-1999; 99US-0134768-
XX 19-MAY-1999; 99US-0134941-
XX 20-MAY-1999; 99US-0135124-
XX 21-MAY-1999; 99US-0135353-
XX 24-MAY-1999; 99US-0135629-
XX 25-MAY-1999; 99US-0136021-
XX 27-MAY-1999; 99US-0136392-
XX 28-MAY-1999; 99US-0136782-
XX 01-JUN-1999; 99US-0137222-
XX 03-JUN-1999; 99US-0137528-
XX 04-JUN-1999; 99US-0137502-
XX 07-JUN-1999; 99US-0137724-
XX 08-JUN-1999; 99US-0138094-
XX 10-JUN-1999; 99US-0138540-
XX 10-JUN-1999; 99US-0138847-
XX 14-JUN-1999; 99US-0139119-
XX 16-JUN-1999; 99US-0139452-
XX 16-JUN-1999; 99US-0139453-
XX 17-JUN-1999; 99US-0139454-
XX 18-JUN-1999; 99US-0139455-
XX 18-JUN-1999; 99US-0139456-
XX 18-JUN-1999; 99US-0139457-
XX 18-JUN-1999; 99US-0139458-
XX 18-JUN-1999; 99US-0139459-
XX 18-JUN-1999; 99US-0139460-
XX 18-JUN-1999; 99US-0139461-
XX 18-JUN-1999; 99US-0139462-
XX 18-JUN-1999; 99US-0139463-
XX 18-JUN-1999; 99US-0139750-
XX 18-JUN-1999; 99US-0139763-

```


PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145493.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 27.9% Score 649; DB 21; Length 437;

Best Local Similarity 35.9%; Pred. No. 8; le-57; Indels 42; Gaps 10;

Matches 165; Conservative 82; Mismatches 177;

OY 3 VFIAFAAALGPLSFSGFALGYSSPAIPSLQRAAPPARLDDAAASWFGAVVTLGAAGGV 62
DB 2 VYLSFVAVGSGFATGSCGYSSPA-----qaalndlsitlaiefsifglitfgamigai 57
OY 63 LGGMLVDNAGRKLSLLCVCPPVAGFAVTTAAQDWMMLGGRLLTGLACGVASLAPVYI 122
DB 58 tsqpladiivgrtkgamvsaefcvvgwlaillfakyyvaldigrletqygmgaftsyyvpilfi 117
OY 123 SEIAYPAVRGLGSCVOLMWWVIGILAYLAGVLEWRMLAVGCVPSIMLLMCPMPT 182
DB 118 aeiakctftrgalctlnqillictgvasvffigrlvkwrvlailqilipcaasfifglffipes 177
OY 183 PRELLTQHRROE-AIALRFLWSEQGWEDDPITGAEOSEFLALR-----OPCIY 230
DB 178 prwlakvgrdtefaealrklrgkkadisee--aaeiqdyietlerlpkamdifqryi 235
OY 231 KPFTIGVSLMAFOQLSGVAAVMFYAFTTFEEAKFKDSSLASVWVCVIOVLFVAVALIMD 290
DB 236 rsvliatfglmvtrfglfgylngicflyssltedagqf-ptrjgmliayavltalnapivd 294

PR 30-APR-1999;	99US-0132407.	PR 21-JUL-1999;	99US-0145088.
PR 04-MAY-1999;	99US-0132484.	PR 22-JUL-1999;	99US-0145085.
PR 05-MAY-1999;	99US-0132485.	PR 22-JUL-1999;	99US-0145087.
PR 06-MAY-1999;	99US-0132486.	PR 22-JUL-1999;	99US-0145089.
PR 06-MAY-1999;	99US-0132487.	PR 22-JUL-1999;	99US-0145192.
PR 07-MAY-1999;	99US-0132863.	PR 23-JUL-1999;	99US-0145145.
PR 11-MAY-1999;	99US-0134256.	PR 23-JUL-1999;	99US-0145218.
PR 14-MAY-1999;	99US-0134219.	PR 26-JUL-1999;	99US-0145224.
PR 14-MAY-1999;	99US-0134221.	PR 27-JUL-1999;	99US-0145276.
PR 14-MAY-1999;	99US-0133370.	PR 27-JUL-1999;	99US-0145913.
PR 18-MAY-1999;	99US-0134768.	PR 27-JUL-1999;	99US-0145918.
PR 19-MAY-1999;	99US-0134941.	PR 27-JUL-1999;	99US-0145919.
PR 20-MAY-1999;	99US-0135124.	PR 28-JUL-1999;	99US-0145951.
PR 21-MAY-1999;	99US-0135353.	PR 02-AUG-1999;	99US-0146386.
PR 24-MAY-1999;	99US-0135629.	PR 02-AUG-1999;	99US-0146388.
PR 25-MAY-1999;	99US-0136021.	PR 02-AUG-1999;	99US-0146389.
PR 27-MAY-1999;	99US-0136392.	PR 03-AUG-1999;	99US-0147038.
PR 28-MAY-1999;	99US-0136782.	PR 04-AUG-1999;	99US-0147204.
PR 01-JUN-1999;	99US-0137222.	PR 04-AUG-1999;	99US-0147302.
PR 03-JUN-1999;	99US-0137528.	PR 05-AUG-1999;	99US-0147192.
PR 04-JUN-1999;	99US-0137502.	PR 05-AUG-1999;	99US-0147260.
PR 07-JUN-1999;	99US-0137724.	PR 06-AUG-1999;	99US-0147303.
PR 08-JUN-1999;	99US-0138094.	PR 06-AUG-1999;	99US-0147416.
PR 10-JUN-1999;	99US-0138540.	PR 09-AUG-1999;	99US-0147493.
PR 10-JUN-1999;	99US-0138847.	PR 09-AUG-1999;	99US-0147935.
PR 14-JUN-1999;	99US-0139119.	PR 11-AUG-1999;	99US-0148171.
PR 16-JUN-1999;	99US-0139452.	PR 11-AUG-1999;	99US-0148319.
PR 16-JUN-1999;	99US-0139455.	PR 12-AUG-1999;	99US-0148341.
PR 17-JUN-1999;	99US-0139452.	PR 13-AUG-1999;	99US-0148365.
PR 18-JUN-1999;	99US-0139454.	PR 13-AUG-1999;	99US-0148684.
PR 18-JUN-1999;	99US-0139455.	PR 16-AUG-1999;	99US-0149368.
PR 18-JUN-1999;	99US-0139456.	PR 17-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139457.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139455.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139460.	PR 20-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139462.	PR 23-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139463.	PR 25-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139750.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151065.
PR 21-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151066.
PR 22-JUN-1999;	99US-0139899.	PR 27-AUG-1999;	99US-0151080.
PR 23-JUN-1999;	99US-0140353.	PR 30-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140354.	PR 31-AUG-1999;	99US-0151438.
PR 24-JUN-1999;	99US-0140695.	PR 01-SEP-1999;	99US-0151930.
PR 28-JUN-1999;	99US-0140823.	PR 07-SEP-1999;	99US-0152363.
PR 29-JUN-1999;	99US-0140991.	PR 10-SEP-1999;	99US-0153070.
PR 30-JUN-1999;	99US-0141287.	PR 13-SEP-1999;	99US-0153758.
PR 01-JUL-1999;	99US-0141842.	PR 15-SEP-1999;	99US-0154019.
PR 01-JUL-1999;	99US-0142154.	PR 16-SEP-1999;	99US-0154039.
PR 02-JUL-1999;	99US-0142055.	PR 20-SEP-1999;	99US-0154779.
PR 06-JUL-1999;	99US-0142390.	PR 22-SEP-1999;	99US-0155139.
PR 08-JUL-1999;	99US-0142803.	PR 23-SEP-1999;	99US-0155486.
PR 09-JUL-1999;	99US-0142920.	PR 24-SEP-1999;	99US-0155659.
PR 12-JUL-1999;	99US-0142977.	PR 28-SEP-1999;	99US-0156458.
PR 13-JUL-1999;	99US-0143542.	PR 29-SEP-1999;	99US-0156596.
PR 14-JUL-1999;	99US-0143622.	PR 04-OCT-1999;	99US-0157117.
PR 15-JUL-1999;	99US-0144005.	PR 05-OCT-1999;	99US-0157753.
PR 16-JUL-1999;	99US-0144085.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144086.	PR 07-OCT-1999;	99US-0158029.
PR 19-JUL-1999;	99US-0144325.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144331.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159331.
PR 21-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159637.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0159584.
			99US-0160741.

PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Query Match 27.9%; Score 649; DB 21; Length 463;

Best Local Similarity 35.9%; Pred. No. 8; Be-57;

Matches 165; Conservative 82; Mismatches 171; Indels 42; Gaps 10;

QY 3 VFLAFAALGSLGSLGSGFALGYSPPALPSIQRAAPPAPRLDDAASWPGAVTILGAAGV 62
DB 28 VYLSFVAVGSAFATGSCAGYSSPA---qaairndlsitlaefslfyslltfamigal 83
QY 63 LCGWLVDRAGRKLSLLCSPFVACFAVITAODVWMLLGGRLTGLACGVASLVAPVYI 122
DB 84 tsgrpiadiygrtgamvssafcvvgyalaitakyyvalidgriatqymgaftsyyvpilfi 143
QY 123 SETAYPAVAGLLGSCVQVLMVVGILLAYLAGVYLEWRMLAVLQCVPPSLMLLMCFMPEY 182
DB 144 aelapktfgrgalitltnqillicgsvsfligtivtrvalialiglipcaasflgtlfipes 203
QY 183 PRFLLTQHRQF-ALALRFLKSGEOWEPPIGAEOSFHLALR-----QPGIT 230
DB 204 pwlakvgrdtfeaalrkrirkkadisee--aaelqdytelrlerlpkakmldlfgrryl 261
QY 231 KPFIIGVSLMAFQOLSGVNAVMFYAETIFEEAKFKDSSLASVVGVOYVLTFAVALIMD 290
DB 262 rsvllafgimvfgqfvgingiclytssifegagf-pltgmilyavilyvitalnapivd 320
QY 291 RAGRLLLVLS--GVVWFSTSAFGAYFKLTOGPGNSSHVAISAPVSAQPYDASVGLAW 348
DB 321 ragrkpplllvsatglivgcllaavsfylkv-----hdmahav-----pv----- 360
QY 349 LAVGSMCLFIAGFVAGWGPPIPMILMSEIFPLHKVATGICVLTNMLMAFLVTKFFSSLM 408
DB 361 lavvgimvyigstsegmgampvwmseifpikkyaggmaltlvnwfgawavsylnflm 420
QY 409 EVLRPGAFWLASFCIFSVLFTLFCVPETKGTLEQITA 448
DB 421 s-wssygtfllyaalnaiaivfaivpetkqktleqiga 459

Search completed: February 13, 2002. 21:50:11
Job time: 6262 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 20:07:13 : Search time 72.65 Seconds
(without alignments)
140,316 Million cell updates/sec

Title: US-09-516-493-7
Sequence: 1 RRVFLAFAAALGPLSFGFA.....CVPETKKTLEQITAFHFGGR 453

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues
Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	571	24.6	494	2	US-09-031-392-5
2	571	24.6	494	4	US-09-299-549-5
3	561.5	24.2	493	2	US-09-031-392-10
4	561.5	24.2	493	4	US-09-299-549-10
5	518.5	22.3	584	2	US-08-928-692-13
6	516	22.2	509	2	US-09-031-392-6
7	516	22.2	509	4	US-09-299-549-6
8	504	21.7	524	2	US-08-928-692-12
9	492	21.2	492	2	US-08-355-844-3
10	492	21.2	492	5	PCT-US95-16126-3
11	480	20.7	534	2	US-09-031-392-4
12	480	20.7	534	4	US-09-299-549-4
13	466.5	20.1	488	2	US-08-928-692-10
14	447.5	19.3	383	2	US-09-031-392-3
15	447.5	19.3	383	4	US-09-299-549-3
16	419.5	18.1	500	2	US-09-031-392-7
17	419.5	18.1	500	4	US-09-299-549-7
18	390	16.8	488	2	US-08-928-692-11
19	378.5	16.3	563	2	US-09-031-392-2
20	378.5	16.3	563	4	US-09-299-549-2
21	259	11.1	555	3	US-08-501-572-3
22	259	11.1	555	3	US-09-040-444-3
23	248	10.7	556	3	US-08-501-572-1
24	248	10.7	556	3	US-08-501-572-1
25	220.5	9.5	553	3	US-09-040-444-2
26	220.5	9.5	553	3	US-09-040-444-2
27	177.5	7.6	455	1	US-08-035-928-2

28	173	7.4	396	2	US-08-850-880-4	21, Appl
29	173	7.4	396	2	US-08-944-916-4	Sequence 22, Appl
30	173	7.4	396	2	US-08-814-877-4	Sequence 21, Appl
31	138	5.9	520	4	US-08-964-127-2	Sequence 14, Appl
32	137	5.9	196	1	US-08-063-552-7	Sequence 17, Appl
33	137	5.9	196	5	PCT-US93-05704-7	Sequence 14, Appl
34	137	5.9	222	1	US-08-035-928-21	Sequence 7, Appl
35	129.5	5.6	337	2	US-08-647-397-2	Sequence 8, Appl
36	117	5.0	194	1	US-08-063-552-8	Sequence 9, Appl
37	117	5.0	144	5	PCT-US93-05704-8	
38	115	4.9	528	2	US-08-403-8520-21	
39	115	4.9	528	3	US-08-510-646B-22	
40	115	4.9	528	4	US-09-231-818-21	
41	110.5	4.8	423	2	US-08-494-907-14	
42	110.5	4.8	423	5	US-08-494-907-14	
43	108.5	4.7	434	2	PCT-US96-10986-14	
44	106	4.6	432	2	US-08-677-049-8	
45	105	4.5	465	1	US-08-471-496-9	

ALIGNMENTS

RESULT 1
US-09-031-392-5
; Sequence 5, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Meng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING GLUTEX AND USFS THEREOF
; NUMBER OF INVENTION: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Melkiohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-031-392-5

Query Match 24.6%, Score 571, DB 2: Length 494;
Best Local Similarity 30.5%, Pred. No. 2.9e-47;
Matches 149; Conservative 94; Mismatches 179; Indels 66; Gaps 10;
OY 3 VFLAFAAALGPLSGFALGYSSPAIFSL-----GRAAPPAPRIIDAAASWFGAV 52
DB 10 LFLFAISATIGSFQGVNTGVINAPPAIKPLNTTLERSSTPPSSVLLTSMISLVAI 69

```

QY 53 VTGGAAGVGLGMLVDRAGRKLSLLCSVPFVAGFAVI---TAADYVMMLLGGRLTLG 109
Db 70 FSVGGMISFSVGLFVFNFRNSMLIVNLAIAGCLMGFKIAESVEMLLIGRLITGL 129
QY 110 AGVASLAVPYISEIAYPAVRGLLSCVQLMNVVGIILAYLAGVNL---EMRMLAVLG 165
Db 130 FCGLCTGVPMYIGLISPTALRGAGFTLNOLGIVIGILVAQIFGLKVLITGEDIMPLLLG 189
QY 166 --CVPSSLMLLMCMPTPRFLTQHRROEAL--ALRFLMSBEGWEDPPIGAQSFH- 220
Db 190 FTILPAIIQCALPFCPESPRFLINRKEEKAKEITLDRLWGTEDVADIDEMKDESRM 249
QY 221 -----LALLRQGIYKPFITIGVSLMAFQOLSGVNAVMFAETIFEAKFKDSSLASY 272
Db 250 SOEKQVTVLELFRAPNYRQPIIISIMLOLSQISINAVFYSTGIFKDAQVQEPVYATI 309
QY 273 VVGVIQVLEFVAALIMDRAGRLLLVLSGVVMVSTSAFCAVFKLQGGNGSSHVAIS 332
Db 310 GAGVNTIFTYVSFVLERAGRRLHLI-----GLGMAFCSIL 348
QY 333 APVSAQPVDAVGLAMLVAGSMCLFIAGFAVGMGPIMLMSLIFPLHVGVATGICVLT 392
Db 349 MTSLLKDNYSWMSFICAILVFVAFELGPGPIPMFTVAELFGGPRPAAVAAGCS 408
QY 393 NMLMAFLVTKFESSIMEVLRPYGAFWL-ASAFCTFSV-----LFTLCVPETKGTLEQ 445
Db 409 NMTSNFLV-----GLFPSATFYLGAIVFIVFVFLVFWVTFEKKVETGRGRFEE 460
QY 446 ITHAFEGR 453
Db 461 ITRAFEGQ 468

```

RESULT 2

```

US-09-299-549-5
: Sequence 5, Application US/09299549
: Patent No. 6136547
: GENERAL INFORMATION:
: APPLICANT: Tartaglia, Louis A.
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: Windows95
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/299,549
: FILING DATE: 26-APR-1999
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/031,392
: FILING DATE: 26-FEB-1998
: ATTORNEY/AGENT INFORMATION:
: NAME: Meiklejohn, Ph.D., Anita L.
: REGISTRATION NUMBER: 35,283
: REFERENCE/DOCKET NUMBER: 07334/072002
: TELEPHONE: 617/542-5070
: TELEFAX: 617/542-8906
: TELEX: 200154
: INFORMATION FOR SEQ. ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 494 amino acids
: TYPE: amino acid

```

```

: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-299-549-5

```

```

Query Match 24.6%; Score 571; DB 4; Length 494;
Best Local Similarity 30.5%; Pred. No. 2,9e-47;
Matches 149; Conservative 94; Mismatches 179; Indels 66; Gaps 10;

```

```

QY 3 VFLAFAAALGPLSEFGALGYSSPAIPLS-----ORAAPPARLDDAANAFGAV 52
Db 10 LITRAIATIGSFQFGVNTGVINAPPAIINKDPLNLTLEERSETPSSVLLTSLMSLVAI 69
QY 53 VTGGAAGVGLGMLVDRAGRKLSLLCSVPFVAGFAVI---TAADYVMMLLGGRLTLG 109
Db 70 FSVGGMISFSVGLFVFNFRNSMLIVNLAIAGCLMGFKIAESVEMLLIGRLITGL 129
QY 110 AGVASLAVPYISEIAYPAVRGLLSCVQLMNVVGIILAYLAGVNL---EMRMLAVLG 165
Db 130 FCGLCTGVPMYIGLISPTALRGAGFTLNOLGIVIGILVAQIFGLKVLITGEDIMPLLLG 189
QY 166 --CVPSSLMLLMCMPTPRFLTQHRROEAL--ALRFLMSBEGWEDPPIGAQSFH- 220
Db 190 FTILPAIIQCALPFCPESPRFLINRKEEKAKEITLDRLWGTEDVADIDEMKDESRM 249
QY 221 -----LALLRQGIYKPFITIGVSLMAFQOLSGVNAVMFAETIFEAKFKDSSLASY 272
Db 250 SOEKQVTVLELFRAPNYRQPIIISIMLOLSQISINAVFYSTGIFKDAQVQEPVYATI 309
QY 273 VVGVIQVLEFVAALIMDRAGRLLLVLSGVVMVSTSAFCAVFKLQGGNGSSHVAIS 332
Db 310 GAGVNTIFTYVSFVLERAGRRLHLI-----GLGMAFCSIL 348
QY 333 APVSAQPVDAVGLAMLVAGSMCLFIAGFAVGMGPIMLMSLIFPLHVGVATGICVLT 392
Db 349 MTSLLKDNYSWMSFICAILVFVAFELGPGPIPMFTVAELFGGPRPAAVAAGCS 408
QY 393 NMLMAFLVTKFESSIMEVLRPYGAFWL-ASAFCTFSV-----LFTLCVPETKGTLEQ 445
Db 409 NMTSNFLV-----GLFPSATFYLGAIVFIVFVFLVFWVTFEKKVETGRGRFEE 460
QY 446 ITHAFEGR 453
Db 461 ITRAFEGQ 468

```

RESULT 3

```

US-09-031-392-10
: Sequence 10, Application US/09031392
: Patent No. 5942398
: GENERAL INFORMATION:
: APPLICANT: Tartaglia, Louis A.
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: Windows95
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/031,392
: FILING DATE: 26-FEB-1998
: ATTORNEY/AGENT INFORMATION:
: NAME: Meiklejohn, Ph.D., Anita L.

```



```

:
: REGISTRATION NUMBER: 35,283
: REFERENCE/DOCKET NUMBER: 07334/072001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617/542-5070
: TELEFAX: 617/542-8906
: TELEX: 200154
:
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 493 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-031-392-10
:
Query Match      24.2%; Score 561.5; DB 2; Length 493;
Best Local Similarity 32.0%; Pred. No. 2.4e-46;
Matches 158; Conservative 87; Mismatches 160; Indels 89; Gaps 14;
:
QY 5 LAFAAALPLSFGFALGYSSPAIPSLQRAAPPAPRLDDAAAMFG-----AVVTIGA- 57
   || |||| :|: || :|: || :|: || :|: || :|: || :|: || :|: || :|: ||
Db 16 LAVLIAALGSFOYGYNLGVIN-----APQKVIKAFYETWLGKXGEXSPVTLTLL 65
:
QY 58 -----AAGVIGGLWLV-----DRAGRKLSLLCSVPFVAG---FAVITTAQDVWMLLG 102
   || |||| :|: || :|: || :|: || :|: || :|: || :|: || :|: || :|: ||
Db 66 WLSLSIFFAVGMIQSGFLVGXIGNRLGRKXAMLVNNVLAIAGGLLMGLAKXKXSPFEMILL 125
:
QY 103 GRLLTGLGAGVASLVAAPYISETIAPPAVAVGILGSCVOLMWVVGILLAYLAGWLE----- 157
   || || |||| :|: || :|: || :|: || :|: || :|: || :|: || :|: || :|: ||
Db 126 GRFTIIGLCYGLSSGVVPMYVGEISPTALRGALGTLMQGLIVIGILLIAOVLG--LDSLGN 183
:
QY 158 ---WRMLAVLIGCVPSLMLLMCFMPETPRLLTQHRROEA---IALRFLWG----- 203
   || || || || || || || || || || || || || || || || || || || || || ||
Db 184 ESLMPLLLGLTGVPALLQLLLLPFCPSRPYLLI-NKNEEARAKKALQRLRGTAIVDSQEV 242
:
QY 204 SEQGWEDPPIGAESFH-LALLRQPGIYKPFIIQVSLMAFOOLSGVNAVAFYAETTFEEA 262
   || || || || || || || || || || || || || || || || || || || || || ||
Db 243 AEMKDESRKMXSEKXSVLELFRSRXYRQPYIAIVLQLSQLSGINAVFYSTSIFEKA 302
:
QY 263 KFKDSSLASVVVGVIOVFTFAVAALIMDRAGRLLLVLSGVVNVFSTSAFGAYFKLTQGG 322
   || || || || || || || || || || || || || || || || || || || || || ||
Db 303 GVGQPYATITAGVYNTVFYVSVFVERAGRRTLHL-----G 341
:
QY 323 PGNSSHVAISAPVSAQPYDASVGLAMLAAGSMCLFTAGFAVGCPIPLMLMSEIFPLHYK 382
   || || || || || || || || || || || || || || || || || || || || || ||
Db 342 LGMAGCAVLMITIALALDQVPMMSVSIIVAFGFVAFVEGPGPIPMFIVALELFSQGR 401
:
QY 383 GVAATGICVLTNNLMAFLVTKFEFSSLMELVRPYGAFWLASAFCTFSVL-----FTLFCVP 436
   || || || || || || || || || || || || || || || || || || || || || ||
Db 402 PAALVAVGFSNMTSNFIVGLLFQYIAELLGPY-----VFIVFAVLLLFIFFTFLKVP 454
:
QY 437 ETGKGTLEQITAHF 450
   || || || || || || || || || || || || || || || || || || || || || ||
Db 455 ETGKGTDEIAAAF 468
:
RESULT 4
US-09-299-549-10
: Sequence 10, Application US/09299549
: Patent No. 6136547
: GENERAL INFORMATION:
: APPLICANT: Tartaglia, Louis A.
: APPLICANT: Meng, Xun
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
: TITLE OF INVENTION: GLUTEX AND USES THEREOF
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804

```

```

:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows95
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/299,549
: FILING DATE: 26-APR-1999
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/031,392
: FILING DATE: 26-FEB-1998
: ATTORNEY/AGENT INFORMATION:
: NAME: Weiklejohn, Ph.D., Anita L.
: REGISTRATION NUMBER: 35,283
: REFERENCE/DOCKET NUMBER: 07334/072002
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617/542-8906
: TELEFAX: 617/542-8906
: TELEX: 200154
:
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 493 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-299-549-10
:
Query Match      24.2%; Score 561.5; DB 4; Length 493;
Best Local Similarity 32.0%; Pred. No. 2.4e-46;
Matches 158; Conservative 87; Mismatches 160; Indels 89; Gaps 14;
:
QY 5 LAFAAALPLSFGFALGYSSPAIPSLQRAAPPAPRLDDAAAMFG-----AVVTIGA- 57
   || |||| :|: || :|: || :|: || :|: || :|: || :|: || :|: || :|: ||
Db 16 LAVLIAALGSFOYGYNLGVIN-----APQKVIKAFYETWLGKXGEXSPVTLTLL 65
:
QY 58 -----AAGVIGGLWLV-----DRAGRKLSLLCSVPFVAG---FAVITTAQDVWMLLG 102
   || || || || || || || || || || || || || || || || || || || || || ||
Db 66 WLSLSIFFAVGMIQSGFLVGXIGNRLGRKXAMLVNNVLAIAGGLLMGLAKXKXSPFEMILL 125
:
QY 103 GRLLTGLGAGVASLVAAPYISETIAPPAVAVGILGSCVOLMWVVGILLAYLAGWLE----- 157
   || || || || || || || || || || || || || || || || || || || || || ||
Db 126 GRFTIIGLCYGLSSGVVPMYVGEISPTALRGALGTLMQGLIVIGILLIAOVLG--LDSLGN 183
:
QY 158 ---WRMLAVLIGCVPSLMLLMCFMPETPRLLTQHRROEA---IALRFLWG----- 203
   || || || || || || || || || || || || || || || || || || || || || ||
Db 184 ESLMPLLLGLTGVPALLQLLLLPFCPSRPYLLI-NKNEEARAKKALQRLRGTAIVDSQEV 242
:
QY 204 SEQGWEDPPIGAESFH-LALLRQPGIYKPFIIQVSLMAFOOLSGVNAVAFYAETTFEEA 262
   || || || || || || || || || || || || || || || || || || || || || ||
Db 243 AEMKDESRKMXSEKXSVLELFRSRXYRQPYIAIVLQLSQLSGINAVFYSTSIFEKA 302
:
QY 263 KFKDSSLASVVVGVIOVFTFAVAALIMDRAGRLLLVLSGVVNVFSTSAFGAYFKLTQGG 322
   || || || || || || || || || || || || || || || || || || || || || ||
Db 303 GVGQPYATITAGVYNTVFYVSVFVERAGRRTLHL-----G 341
:
QY 323 PGNSSHVAISAPVSAQPYDASVGLAMLAAGSMCLFTAGFAVGCPIPLMLMSEIFPLHYK 382
   || || || || || || || || || || || || || || || || || || || || || ||
Db 342 LGMAGCAVLMITIALALDQVPMMSVSIIVAFGFVAFVEGPGPIPMFIVALELFSQGR 401
:
QY 383 GVAATGICVLTNNLMAFLVTKFEFSSLMELVRPYGAFWLASAFCTFSVL-----FTLFCVP 436
   || || || || || || || || || || || || || || || || || || || || || ||
Db 402 PAALVAVGFSNMTSNFIVGLLFQYIAELLGPY-----VFIVFAVLLLFIFFTFLKVP 454
:
QY 437 ETGKGTLEQITAHF 450
   || || || || || || || || || || || || || || || || || || || || || ||
Db 455 ETGKGTDEIAAAF 468
:
RESULT 5
US-08-928-692-13
: Sequence 13, Application US/08928692
: Patent No. 5958727

```

GENERAL INFORMATION:

APPLICANT: Brody, Howard
 APPLICANT: Yaver, Deborah S.
 APPLICANT: Lamsa, Michael
 APPLICANT: Hansen, Kim
 TITLE OF INVENTION: Methods for Modifying the Production of
 TITLE OF INVENTION: a Polypeptide
 NUMBER OF SEQUENCES: 80
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 59587270 No. 5958727disk of No. 5958727th America, Inc.
 STREET: 405 Lexington Avenue
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10174
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,692
 FILING DATE: 12-SEPT-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 4944,200-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-867-0123
 TELEFAX: 212-878-9655
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 584 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: No. 5958727e
 US-08-928-692-13

Query Match 22.3%; Score 518.5; DB 2; Length 584;
 Best Local Similarity 28.6%; Pred. No. 4,7e-42;

Matches 138; Conservative 81; Mismatches 200; Indels 63; Gaps 10;

3 VFLLAFAAALPLSPFGALGYSSPAIRSLORAPPAARLDDAAASV-----FGAVVTGGA 57
 86 IITLTVASISGMEFGIDTGYISSALISI-----GTDLHKVLTGKEIYVAATSLGA 139
 58 AAGVILGWLVDRAKRLSLILGSPFVAVFAVITAADVWMLGRLTLGLACGVASLY 117
 140 LITSIFAGTAADIFGRKRLMGSNLMFVIGAILQVSAHTFWMQAVGRLLMGFGVIGSLI 199
 118 APYISEIAYPAVRGLLGLSCVOLMNVVGLILAVLAGVLE-----WRWLAVLGCVPPSLM 172
 200 APFISEIAPKMKIRGLTYINSIMLIGSOLVAVGCGAGLNVNNGRLLVGLSLIPTAVO 259
 173 LILMCPMPFPLLTQ---HRRQEAIALRFLMGSEOGMEDPPICAEOSFHLALLRQ--- 226
 260 FTLICFLPTPRYVWKGLARATEYLKRSYTDTSBE-----IIRKVEELVLTNOSIP 313
 227 -----PGIYKPFILIGVSLMAFOOLSGVNAVMFAETIPEAKEKDS 268
 314 GKNVEKVMNTIKELHTVSNLRALITGGGLOAIQOFTGWNLSLMFSGTIEFTVGFKMS 373
 269 LASVVGVVYVFTVAALIMDRAGR--LLVLISGVVMFSTSAFGAYFKLTGGPGNS 326
 374 AVSIIIVSGNFTFLVAFPSIDKIRKTLILGLPGMTMALVVCISIAHF----- 423
 327 SHVAISAPVDAQPDASVGLAV--LAVGSMCLFIFAGFVAGMPFVWMLMSELPFHYKGV 384
 424 --LGIKFQGAIVAVVVSFGSSGMIYIIVFIIVFAAFYALGIGTYPW--QOSLFPQNVKGI 480

385 ARGICVLTNMIAFLVTKFESSIMEVLRPYGAFMLSAFCIFSVLETLFCVPEKTKLE 444
 481 GTSVATATNMAGSLVIASTFLTMQNTIPACTFAFAAGLSCLSLTFCYFCELSGLELE 540
 445 QI 446
 541 EV 542

RESULT 6
 US-09-031-392-6

Sequence 6, Application US/09031392
 Patent No. 5942398

GENERAL INFORMATION:

APPLICANT: Tartaglia, Louis A.

APPLICANT: Weng, Xun

TITLE OF INVENTION: NUCLEIC ACID MOLECULES
 TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/031,392

FILING DATE: 26-FEB-1998

ATTORNEY/AGENT INFORMATION:

NAME: Meiklejohn, Ph.D., Anita L.

REGISTRATION NUMBER: 35,283

REFERENCE/DOCKET NUMBER: 07334/072001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 509 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-031-392-6

Query Match 22.2%; Score 516; DB 2; Length 509;

Best Local Similarity 30.8%; Pred. No. 6,7e-42;

Matches 155; Conservative 83; Mismatches 163; Indels 102; Gaps 16;

5 LAFAAALPLSPFGA-----LCYSSPAIRSLORAPPAARLDDAA 45
 26 LAVFSAGVILGQFQVINGVIAAPQKVIQGSYNETWLRGQEPGS---STP----- 74
 46 ASWFGAVVTGGA-----AAGVILGWLVDRAKRLSLILGSPFVAVFAVITAADVWMLGRLTLGLACGVASLY 152
 75 -----GTLTLMALSAIFSVGGMISFLIGIISQWLGRKRAMLVNVLAVLGGSLMGLAN 130
 93 AADVWMLGRLTLGLACGVASLYAVYITSEIAYPAVRGLLGLSCVOLMNVVGLILAVLAGVLE-----WRWLAVLGCVPPSLM 172
 131 AAASYEMLLIGRLIGVSGITGLVPMVYGETAPHLRGALGTMLQALVIGILTAQVL 190
 153 GVWLE-----WRWLAVLGCVPPSLMILMCPMPFPLLTQ---HRRQEAIALRFLMGSEOGMEDPPICAEOSFHLALLRQ--- 226
 191 G--LESLLGASLMPPLILGLTFLPALLOLVLLPFCSPESPRYLITLIONLEGPAKSLKRLT 248
 202 -MGSEOG---WRDPPIGAMQSFHLALLRQPGT---YKPFIIIGVSLMAFOOLSGVNAVMFAETIPEAKEKDS 268

us-09-516-493-7.rai

Page 6

Query Match	21.7%	Score 504	DB 2	Length 524
Best Local Similarity	28.0%	Pred. No. 1e-40		
Matches 146	Conservative 87	Mismatches 172	Indels 116	Gaps 17

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

1 CLASSIFICATION: 436
2
3 ATTORNEY/AGENT INFORMATION:
4 NAME: Tang, Henry Y.S.
5 REGISTRATION NUMBER: 29,705
6 REFERENCE/DOCKET NUMBER: A29927-50/29910
7
8 TELECOMMUNICATION INFORMATION:
9 TELEPHONE: 212-408-2586
10 TELEFAX: 212-765-2519
11 INFORMATION FOR SEQ ID NO: 3:
12 SEQUENCE CHARACTERISTICS:
13 LENGTH: 492 amino acids
14 TYPE: amino acid
15 STRANDEDNESS: single
16 TOPOLOGY: linear
17 MOLECULE TYPE: peptide
18 ORIGINAL SOURCE:
19 ORGANISM: Human
20
21 FEATURE:
22 NAME/KEY: Peptide
23 LOCATION: 1..492
24 OTHER INFORMATION:
25 OTHER INFORMATION: Facilitative glucose transporter
26 OTHER INFORMATION: Glut1 protein
27
28 US-08-355-844 -3

```

Query Match	21.2%	Score 492:	DB 2:	Length 492:
Best local Similarity	28.8%	Pred. No. 1,4e-39;		
Matches 145;	Conservative 90;	Mismatches 167;	Indels 102;	Gaps 15
QY	2	RVELAAFAAALGPIISGFALG-----	YSSPAIPISQRAAPAPRI	41
Db	11	RLMLAVGAVLGLSGLOGYNTGVINAPQKVEEPTYNQTNWHRGEEILPT-----		59
QY	42	DDAAASFCGAVVTLLGAAGVLC-----	GWLVDRAGRKLLSLICS-VPVA-----	GEAVIT 92
Db	60	TLTTLMSLSVAIF--SVGGMISFSVGLFVNFRGSRNSMLMNNLLAFSAVILMGSKL-		115
QY	93	AADVVMILGGRLTLGAGVASLAPVYISLAIYAVAGLIGSCQVLMMVVCILLATYA		152
Db	116	GKSEFMILLGRITLIDYGVGLTGEVPMVGEVSPPAFGALGTLHQLGIVGILLINQF		174
QY	153	GWVLE-----WRMLAVIGCVPSPLMLILMCMFETPRFLLTOHRQEATA-----	LRFL	201
Db	175	G-LDSIMGNKDLMLPLLSITIFPALLOCIYIPFCGESRFLIT-NRNEENRAKSYLKL		231
QY	202	WCSQGWEDPPIGAEDSFH-----	LALLRQGIYKPFITIGVSLMAFOOLSGVNAVM	252
Db	232	RGTAIDVTHLQEMKEESRQMRREKKVITLLEPRSPAYRQPIILAVILQISQOLSGINAVF		291
QY	253	FAEITLFEFAKKKDDSLASVVVYQVQLFTVAALIMDRAGRLLVLISGVYMFSTSAF		312
Db	292	YISTLFEFAAGVQDPYVAITIGGIYNTAFTVSLVEERAGRTTHLIL-----		339
QY	313	GAYFKLLTGGPGNSSHVAISAPVASQPVASGLMMLAVNGSCLEFIAGVAGMGPIPML		372
Db	340	-----GLAGMGAGQALLMILMIALALLEQLPWSMYLSLVAIFGVAAFEVGGPIIPFI		390
QY	373	MSLIPPLHVKYAGICITVLTNMLMAFLVTKFSSLMELVLRPYGAFMLASAPCITVLEPTL		432
Db	391	VAELISQGPRAALIVAAFSNWTSNFIVGMCFQYVEOLGCPY-----	VFIIIFVLLVL	443
QY	433	FC-----VPETKCKTLEQITAHF		450
Db	444	FFIRITFVVPETKGTDEIASGF		467

```

RESULT 10
PCT-US95-16126-3
; Sequence 3, Application PC/TUS9516126
; GENERAL INFORMATION:
; APPLICANT: Fischbarg, Jorge
; APPLICANT: Czegledy, Ferenc
; APPLICANT: Isetovich, Pavel

```

APPLICANT: Li, Jun
APPLICANT: Cheung, Min
TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN
TITLE OF INVENTION: STRUCTURE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16126
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/355,844
FILING DATE: 14-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Tang, Henry Y.S.
REGISTRATION NUMBER: 29,705
REFERENCE/DOCKET NUMBER: A2927-50/29910
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2586
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 492 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Human
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..492
OTHER INFORMATION: Facilitative glucose transporter
OTHER INFORMATION: Glut1 protein
PCT-US95-16126-3

Query Match 21.2%; Score 492; DB 5; Length 492;
Best Local Similarity 28.8%; Pred. No. 1.4e-39;
Matches 145; Conservative 90; Mismatches 167; Indels 102; Gaps 15;

QY 2 RYFLAFAALGRLSGFALG-----YSSPAIFSLORAAPPRL 41
DB 11 RLMLAAGAVLSTLQGYNGVINAPOKVIIEFYNOTVHRHGRSILPT----- 59
QY 42 DDAASWFGAVVTILGAAGCVLG---GMLVDRAGRLSLILCS-VPEVA---GPAVIT 92
DB 60 -LITLIMSVAIF--SVGGMISFSVGLFVNFRGRNSMLMLMLLAFVSIVLMGFSL- 115
QY 93 AADVWMLGRLTLGLAGVASLVAIVYISLAVPAVRGLLSCVOIMVAVGTLILAVLA 152
DB 116 -GKSFEMILIRITIGVGLTGFVPMVYGEVSPAFRGALGLHGLGVIGLILNOVF 174
QY 153 GWVLE-----WRWLAIVGCVPPSLMLLMCFMPTPFFLLTOHROAIA---LRFL 201
DB 175 G--LDSIMKNKDLMLPLSTIFPALLOCIIVIPFCSPSPFLLI-NNNEENRAKSVLTKL 231
QY 202 MCESEQWEPPIGAEQSFH-----LALLROPGIKPFTIGVSLMAROOLSGVAVM 252
DB 232 RGTADVTHLOEKKEESROMMRKKVITILEFRSPAVRQIILAVVLOLSQSLGINAVF 291
QY 253 FYAETIFEEAKFRDSSLASVVGVIQVLFATAVALIMDRAGRLLILVSGVVMVFTSAF 312

DB 292 YSTSIFFKAGYQOPVYATIGSIVNTAFTVVSLEFVERAGRRLTLI----- 339
QY 313 GAYFKLTGGPENSASHVIAISAPVSAQPVDAVSGLMLAVGSCLPFAGVAGPIPWLL 372
DB 340 -----GLAGMAGQALMTLALALAEQLPFWMLSTLVAIFGVAFEEVGGPIPWFI 390
QY 373 MSEIFPLHKVATGICVLTNMLMAFLVTKFSSIMENVLRPGAFMLASAFICFSVLFLL 432
DB 391 VAELESQGRPAALVAVGFSNMTSNFIVGMCFOYVQOLGPI-----VFIIFTVLVL 443
QY 433 FC-----VPETKGTLEQITAHF 450
DB 444 FEIRTFKVPETKGRTPDEIASGF 467

RESULT 11
US-09-031-392-4
Sequence 4, Application US/09031392
Patent No. 5942398
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
APPLICANT: Weng, Xun
TITLE OF INVENTION: NUCLEIC ACID MOLECULES
TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,392
FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Melkiojohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 534 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-031-392-4

Query Match 20.7%; Score 480; DB 2; Length 534;
Best Local Similarity 28.0%; Pred. No. 2.2e-38;
Matches 149; Conservative 91; Mismatches 168; Indels 124; Gaps 18;

QY 5 LAAPFAALGSLFSGALG-YSSP-----AIPSLORA----- 35
DB 19 LSVFPAVIGFPOYGYSLCVINAPKVIIEAHNGRMGLALPMVHRHATNTSRDNATTIVTIG 78
QY 36 -----PPAPRLDDAAS-----WFGAVVTILGAAGCVLG---GMLVDRAGRL 75
DB 79 TEAMSSSEGTILAPSAQFEDPIYSPHILIMYWSLSVMAVAGMWSFTYGMWIGDLGRVK 138
QY 76 SLICSVFVAGFAVITAAQ---DVMMLLGRLLTGLACGVASLVAIVYISLAVPAVRG 132
DB 139 AMLVNVNLSIAGNLLMGLAKGSPSHIITIIAGRAITGLYCGLSGLVPMYVSEVSPALRG 198

Query Match	20.7%;	Score 480;	DB 4;	Length 534;
Best Local Similarity	28.0%;	Pred. No. 2.2e-38;		

US-08-928-699-10
Sequence 10, Application US/08928692
Patent No. 5958727
GENERAL INFORMATION:
APPLICANT: Brody, Howard
APPLICANT: Yaver, Deborah S.
APPLICANT: Lamsa, Michael
APPLICANT: Hansen, Kim
TITLE OF INVENTION: Methods for Modifying the Production of
NUMBER OF SEQUENCES: 80
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5958727/o No. 5958727/disk of No. 5958727/1h America, Inc
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,692
FILING DATE: 12-SEPT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4944.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 20:10:58 ; Search time 79.58 Seconds

(without alignments)
433.615 Million cell updates/sec

Title: US-09-516-493-7

Perfect score: 2324

Sequence: 1 RRVFLAFAAALGPLSGFA.....CVPEKCKTLEDTAHEGR 453

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:
2: PIR2:
3: PIR3:
4: PIR4:

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	725.5	31.2	487	2	E96782 hypothetical prote
2	717.5	30.9	490	2	T14545 probable sugar tra
3	601.5	25.9	461	2	D70073 metbolute transpo
4	596	25.6	472	2	S47089 arabinoase-proton s
5	593	25.5	464	2	F65079 galactose-proton s
6	589	25.3	464	2	F85951 galactose-proton s
7	589	25.3	472	2	B26430 l-arabinose isomer
8	589	25.3	472	2	E85936 hypothetical prote
9	577.5	24.8	521	2	G84864 probable membrane
10	574	24.7	575	2	T43400 myo-inositol trans
11	572	24.6	496	2	T52132 probable sugar tra
12	569	24.5	493	2	S38981 glucose transport
13	566	24.4	493	2	A41751 glucose transport
14	565	24.3	496	2	A31986 glucose transport
15	555.5	23.9	580	2	D86426 hypothetical prote
16	550	23.7	457	2	E70070 metbolute transpo
17	545	23.5	464	2	F69587 l-arabinose transp
18	542	23.3	560	2	T51485 sugar transporter-
19	536	23.1	461	2	G85059 probable sugar tra
20	535.5	23.0	491	2	A26430 xylose transport p
21	535.5	23.0	491	2	B86096 xylose transport p
22	535	23.0	522	2	A31556 glucose transport
23	534.5	23.0	523	2	S25015 monosaccharide tra
24	534	23.0	526	2	T01853 probable hexose tr
25	532.5	22.9	523	2	S06920 glucose transport
26	530.5	22.8	534	2	S38435 hexose transport p
27	529	22.8	468	2	S10014 glucose transport
28	525	22.6	496	2	A41264 glucose transport
29	525	22.6	507	2	T01844 probable sugar tra

30	523.5	22.5	457	2	H85059 probable sugar tra
31	523.5	22.5	514	2	S25009 monosaccharide tra
32	523.5	22.5	522	2	E86246 glucose transport
33	521.5	22.4	516	2	T12199 monosaccharid tran
34	521.5	22.4	612	2	B40538 myo-inositol trans
35	520.5	22.4	522	2	S12042 glucose transport
36	519.5	22.4	547	2	C84593 probable sugar tra
37	518.5	22.3	584	2	S69555 myo-inositol trans
38	517	22.2	509	2	A32101 glucose transport
39	517	22.2	534	2	S14144 hexose transport p
40	516	22.2	509	2	A33801 muscle-fat glucose
41	514.5	22.1	508	2	T05156 probable glucose t
42	513	22.1	492	2	S09705 glucose transport
43	512.5	22.1	523	2	T10122 hexose transport p
44	511	22.0	557	2	T38125 myo-inositol trans
45	509	21.9	492	2	I45902 glucose transporte

ALIGNMENTS

RESULT 1
E96782
hypothetical protein F22H5.6 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: E96782
R:Theologos, A.; Ecker, J. R.; Palm, C. J.; Federspiel, N. A.; Kaul, S.; White, O.; Alon
Chin, C. W.; Chung, M. K.; Conn, L.; Conway, A. B.; Conway, A. R.; Creasy, T. H.; Dewar,
ansen, N. F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J. L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C. A.; Li, J. H.; Li, Y.; Lin, X.; Liu, S. X.; Liu, Z. A.; Lueros, J. S.; Maitl, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S. L.; Schwartz, J. R.; Shin, P.; Southwick, A. M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C. M.; Venter, J. C.; Davis, R. W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; M0ID:21016719
A:Accession: E96782
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1487 <STO>
A:Cross-references: GI:AE005173; NID:q10092276; PIDN:AA012689.1; GSPDB:GN00141
C:Genetics:
A:Gene: F22H5.6
A:Map position: 1
C:Superfamily: glucose transport protein

Query Match 31.2% Score 725.5; DB 2; Length 487;
Best Local Similarity 37.5%; Pred. No. 2.2e-47;
Matches 173; Conservative 77; Mismatches 176; Indels 35; Gaps 7;

QY	3	VFLAFAAALGPLSGFALGYSSPAIRSLQRAAPARLDDAASWFGAVYTTGCAAGCV	62
		: : : :	
DB	48	VLACVLLVALGPIQFGFTCGSSPT---QAIRKIDGLITVSEVSFGSSNWGAMGAI	103
QY	63	LGGLVLRAGCKSLILCSVPVAGFAVITRAODVMMLGLGRLLTGACGASVAVAPYI	122
		: : : : : :	
DB	104	ASGOIAEYIGKRGSLMTAIPNITIGWLCTISPAKDTSLYMKRLLGEGVGVIISTVPIYI	163
QY	123	SEIAYPAVRGLGSCVOLIMVAVGILLAVYLAGWVLEWRMLAVGCVPSLMLLMCFMPE	182
		: : : : :	
DB	164	AEIAPQMMRGISGVNDSVYIGIMLVILGLFVPMRILAVLGLTPTLLIPGLFPIES	223
QY	183	PRFLITTHRRQE-AIALRFLMGSEQGWEDDPICAEQ-----SFLIALLRQPCI	229
		: : : : : : : : :	
DB	224	PRWLAKMGMTDEFTSLQVLKGFET---TDITVEVNETIKRSVASTKRNFVFDLKRRY	280
QY	230	YKPRITGVSLMAPOOLSGVNAVMEFYAEITFEFAFKFSSLASVVGIVGYVFTVAALIM	289
		: : : : : : : : :	
DB	281	YFPLMVGILVLVLOOLGIMGVILFTYSTIFESAGVTSNMAFTGVGAIVAVATISWLV	340

QY 290 DRAGRLLLVLSGVVWFSTAFGAYFKLTGGPGNSHVAISAPVSAQPVDAVGLAML 349
 Db 341 DKARGRLLLTSSVGMTLSLVAAYLKE-----PSPDS---DKMSWLSIL 386
 QY 350 AVSGMCLFAGFVANGKPIWMLSEFLPHVGVATGICVLTNMLMAFLVTKFSSIME 409
 Db 387 SVGVAVVAVVFFSLGMPIMLSEILPVNIKGLASITANLWFWFSWILTM-ANLL 445
 QY 410 VLRRPAGFWLASAFICFVSFTLFCVPEPKTKLEQTIAHF 450
 Db 446 AMSSGGFTLXLGVCAFTVYFVILWPEIKTKLLEQLSLF 486

RESULT 2
 T14545
 probable sugar transporter protein - bee
 C:Species: Beta vulgaris (bee)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 C:Accession: T14545
 R:Chou, T.J.; Bush, D.R.
 Plant. Physiol. 110, 511-520, 1996
 A:Title: Molecular cloning, immunochemical localization to the vacuole, and expression
 A:Reference number: 218131; MUID:96551183
 A:Accession: T14545
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-490 <CHI>
 A:Cross-references: EMBL:043629; NID:g1209755; PIDN:AA53155.1; PID:g1209756
 A:Experimental source: tonoplast
 C:Superfamily: glucose transport protein
 C:Keywords: transmembrane protein

Query Match 30.9%; Score 717.5; DB 2; Length 490;
 Best Local Similarity 36.1%; Pred. No. 8.7e-47;
 Matches 167; Conservative 82; Mismatches 175; Indels 39; Gaps 6;

QY 3 VFLAFAALGRLSPFGALGYSSPAIPSLORAPAPRLDDAAMFGAVVTLGAAGV 62
 Db 50 VLACVLIVAGLGIQGFAGYSSP-----QSATNELGLSAEYSEFGLSNVGMVGA 105
 QY 63 LGQWLVDRAGRKLLSLSCVPFVAGFAVITTAADVWMLLGRLLTGLAGVAVLPYI 122
 Db 106 ASQGVSEYIGRKKSLMIAITNIGMLAISRAKDSFLYGRMLEGFGVGLISTVPIYI 165
 QY 123 SEIAPVAVRGLLSCVQLMNVVGLILAVLWLEWMLAVIGVPSIMLLCMFPER 182
 Db 166 SEIAPQNLGALGVNQLSVTIGIMLSYMLGLEVPWRILAVILGILPCLILGLFIPES 225
 QY 183 PRFLTQHRROE-AIALRFLWGSQGMEDPPIGAE-----QSFHLALLRQP 227
 Db 226 PRWLAKXGMEEFEVSLQVLCF-----DIDISLEVNEIKRVSASSKRTTIRPAELROR 280
 QY 228 GIYKPFITIGVSLMAFQOLSGVNAVVFYAETTFEAKFKDSSLASVVGVIOVLETAVAL 287
 Db 281 RYWLPLMIGNLLITLQOLSGINGVLFYSTIFKEAGVTSSNAARFGCAQAVIATVTTW 340
 QY 288 IMDRAGRLLLVLSGVVWFSTAFGAYFKLTGGPGNSHVAISAPVSAQPVDAVGLA 347
 Db 341 LVDSGRRLILVSSSGMTLSLVAAMSFLEKENVSDSTWYSV-----FS 386
 QY 348 WLAVGSMCLFIAGFVANGKPIWMLSEFLPHVGVATGICVLTNMLMAFLVTKFSS 407
 Db 387 ILVAVGVAVVFFSLGICAIPIIMISEILPINKGLASITANLWFWFAMIVTMT-ANI 445
 QY 408 MEVLRRPAGFWLASAFICFVSFTLFCVPEPKTKLEQTIAHF 450
 Db 446 MLSWNSGTFISYVAVFAFVAVIWPETKRTLEIQMSF 488

RESULT 3
 D70073
 metabolite transport protein homolog yxoc - Bacillus subtilis

C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C:Accession: D70073
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.;
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funo, S.; Galizzi, A.; Gal
 lech, J.; Harwood, C.R.; Henaut, A.; Hillbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
 Y. M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
 Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scant
 A:Authors: Schleich, S.; Schwoerer, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
 T.; Winters, P.; Wipat, A.; Yamamoto, T.; Terstara, P.; Toignon, A.; Tosato, V.; Uchiya
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
 A:Reference number: A69580; MUID:98044033
 A:Accession: D70073
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-461 <KUN>
 A:Cross-references: GB:299124; GB:AL009126; NID:g2636442; PIDN:CAB16017.1; PID:g26365
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yxoc
 C:Superfamily: glucose transport protein

Query Match 25.9%; Score 601.5; DB 2; Length 461;
 Best Local Similarity 31.1%; Pred. No. 4.6e-38;
 Matches 146; Conservative 85; Mismatches 191; Indels 47; Gaps 7;

QY 2 RVLFAFAALGRLSPFGALGYSSPAIPSLORAPAPRLDDAAMFGAVVTLGAAGV 61
 Db 6 RKMVIFFGALGLLYGIDGVCISGALLFINOPLTLLEGVY-----MLLGLAFGS 61
 QY 62 VLGWLVDRAGRKLLSLSCVPFVAGFAVITTAADVWMLLGRLLTGLAGVAVLPYI 121
 Db 62 ALSTGSDRMGRKRRVNVLSLIFIIIGALACAFSGTIGMLIASRVILGLAVGSGFALPVY 121
 QY 122 ISEIYPAVAVRGLLSCVQLMNVVGLILAVLWLEWMLAVIGVPSIMLLCMFPER 177
 Db 122 LSEMAPTKIGTIGTMMNMLIVTIGILATVNLFTFPEAKRMVWGLAAVPAVLLIGIA 181
 QY 178 FMEPTPELLTQHRROEAL-----RFLWSEQGMEDPPIGAEQSFHLALRQ 226
 Db 182 FMEPSRPLVKKRSEERARRINMITHDPKDIEMELAEKMGAE-----KKTTLGLVKA 236
 QY 227 PGYKPFITIGVSLMAFQOLSGVNAVVFYAETTFEAKF--KSSLASVVGVIOVLETAVAL 284
 Db 237 KWIIRPMLITVGLAIHQAVGINTVYVAFITFKAGLGTSAISALGTWIGILNVTMCT 296
 QY 285 AALIMDRAGRLLLVLSGVVWFSTAFGAYFKLTGGPGNSHVAISAPVSAQPVDAVGLA 344
 Db 297 AMILIDRVGRKLLINGSVITLSLALSGVL-LTIG-----LSA 335
 QY 345 GLAWLAVGSMCLFIAGFVANGKPIWMLSEFLPHVGVATGICVLTNMLMAFLVTKF 404
 Db 336 STRAMVVELVGLYIVYQATWGPVVVWLMPELPPSKARAGAFITLVAANLIVSLVF 395
 QY 405 SSLMEVLRRPAGFWLASAFICFVSFTLFCVPEPKTKLEQTIAHF 453
 Db 396 PLMLSMGLAMVPMVFSVCLLSFFAFYWPETKGSLEIEASLKKR 444

RESULT 4
 S47089
 arabinose-proton symporter - Klebsiella oxytoca
 C:Species: Klebsiella oxytoca
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
 C:Accession: S47089

R:Shatwell, K.F.; Charalambous, B.M.; McDonald, T.P.; Henderson, P.J.F.
submitted to the EMBL Data Library, June 1994
A:Description: The nucleotide sequence of the gene *araE* for arabinose-proton symport in
A:Reference number: S47089
A:Accession: S47089
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-472 <SHA>
A:Cross-references: EMBL:X79598; NID:q498919; PIDN:CA56110.1; PID:q498920
C:Superfamily: glucose transport protein
:Keywords: arabinose transport

Query Match	25.6%	Score	596;	DB 2;	Length	472;			
Best Local Similarity	33.7%	Pred. No.	1.2e-37;						
Matches	153;	Conservative	84;	Mismatches	181;	Indels	36;	Gaps	9

QY	7	AFAPAAAGLSTGSPALGSSSPAI	PSLQRAAPAPAPRLDDDAAS	WFGAVNTLGAAGCVLGG	66	
		:	:	:	:	
Db	26	SIAAAVAGLLFGDLGIVAGALP	FTTDHFVLSRRQF----	WVSSMILGAAIGALFNGW	81	
QY	67	LVDRAGRRLSLLLCSVPVAGV	FAVITTAADYMMILGRLTGL	LACGVASLVAAPYISEIA	126	
		:	:	:	:	
Db	82	LSFRLGRKYSILMGCAVLE	SVGSAFATSVEMLLVARIV	LGVAAIGASTAPILYSEMA	141	
QY	127	YPAVAGLLGSCVOLMNVVVG	ILLATYLAGVLEW--	RWLAVLG--	CYPPSLMLLMCFMPT	182
		:	:	:	:	
Db	142	SENRNGKMTSMYOLMNTV	TGIVMAFLSDTAFFSGNWR	MALGVALLPAVALLIVL	FPLNPS	201
QY	183	PRFLTLQRRQEA--	IALRFILMGSEQGMEDPP	IGAROSFHL-----	ALLR-QPGIYKPELI	235
		:	:	:	:	
Db	202	PRMLAEKGRHVEAEFVLE	MLRDTSEKARDELNET	RESLKLQGGVALR	KVNNNVRAVFL	261
QY	236	GVSILMAFOQLSGVNAVMT	YAEETIFEBAKF---	KDSSLASVYVGTQVLT	VAVALIMBRA	292
		:	:	:	:	
Db	262	GMLLOAMQOFTGMNIT	MYAPRIFFMAGFTTEQDM	VAITLVYGLTFEATPI	IAVFTVDKA	321
QY	293	GRRLILVSTGVVWFSTAF	GAFFKLTGGCGPNSSHVA	ISAPVSAQPDVAS	GLAMLVAG	352
		:	:	:	:	
Db	322	GRRRLAKIGESVYMAL	IGTLVLG--	YCLMOPDMGT-----	ASSGLSWLSVG	363
QY	353	SMCLTEFAFVAGWGP	IPWLLMSEIFPLHVKGVAT	ICVLTWMLMAFLYK	EFSSLMEEYLK	412
		:	:	:	:	
Db	364	MTMMCIAGYASAA	PVVYVILCSEIQPLKCR	DGICSTTTMWSMMIL	IGATFLTLDAIG	423
QY	413	PYGAFWLASAFCIFS	VLTFLFCVPETTKCTLE	QI	446	
		:	:	:	:	
Db	424	AAGTFMILTALNVAF	IGVTFWMLIPETKNV	TLEHI	457	

RESULT 5
F65079
galactose-proton symport (galactose transporter) - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence-revision 17-Sep-1997 #text-change 24-Sep-1999
C:Accession: F65079
R:Bialtner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co-
.A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: F65079
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1464 <BLAST>
A:Cross-references: GB:AE000377; GB:U00096; NTD:q2367178; PIND:AACT5980.1; PID:q1789312
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: galP
A:Superfamily: glucose transport protein

Query Match 25.58; Score 593; DB 2; Length 464

Best Local Similarity 32.68; Pred. No. 2e-37;
Matches 156; Conservative 83; Mismatches 173; Indels 66; Gaps 12;

Oy 4 FLAARAAALGPLSPCFALAGYSSPAIPSLQRAAPRAPRPRIDDAASMSFGAVVTLIAAGAGVL 63
 Db 16 FVFCVPLAAGLGLFGLDGLGVLAGALPFTADEF---QITSTHQEVVSSSMFGAAVAVG 71
 Oy 64 GCMVLDRAGRKSLILLCVPPFAGFAVITAADQVMMLLGGRLTGLCGVAISLAAPVYS 123
 Db 72 SGWLSTFKGRKSKSLTGAILFVYAGSLFSAAPANVEVLLSRVLLGLAVGASTAPRLYLS 131
 Oy 124 EIAVPAVEGLDSCVQLMWWVVGILLAYLA---GWLVEPMYLAVLGCVPSLMLLMCFM 179
 Db 132 ELAPRKIGSMISMLQMLITIGLAVLSDTAFSTGTGMRRMLGVIIITPALLLLIGVFLL 191
 Oy 180 PETPFRLLTQHR--RQEAIALRF-----LMSBOGNEEDPFGAEDSFH 220
 Db 192 PDSRPMFAAKRFVDAEVLRLRDTSAEAKRELDETRSLQVQKSG----- 239
 Oy 221 LALLRPGIYKPFITGVSLMAFOOLSGVNAVMTAETIFEDAKFKDSS---LASVYGV 276
 Db 240 -ALFKENSNFRAVFLGVLLOVMQFTGMNIMYAPRIFELAGYNTTEDMGCVITVGL 298
 Oy 277 IOVLETAVALINDAGRRLLLVLSGVYVWVSTSAFGAKFYKLTQGGPONSNNHVALISAPVS 336
 Db 299 TNNVLTFTFIAIGLVDWGRKPTLTJGELV---AAG-----MGVLCTMMHIGHSIP-S 346
 Oy 337 AOPVDAVGLAMLAAGSMCLFAGFAVGMGPRMLMSEIFPLHVKVATGICVLTNMLM 396
 Db 347 AO-----YFAIAMLMTFVGFMAAGPLIWLVCSEIOPKLRGDPGTCISATNMIA 397
 Oy 397 AFLVTKFFSSLMEMVIRPYGAFMLASAPCIFSVLFTLFCVPEPTKGTLYRQIITAH-FEGR 453
 Db 398 NMIVGATFELTMLNTLGNMNTFMYVAALNVLFLLTJMLVPEPKHVSLEHIERNLMKGR 455

RESULT 6
 F85951
 galactose-proton symport of transport system [imported] - Escherichia coli (strain O157:H7)
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
 C:Accession: F85951
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
 11ler, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11306551
 A:Accession: F85951
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-464 <STO>
 A:Cross-references: GB:AE005174; NID:q12517486; PIDN:AA658074.1; GSPDB:GN00145; UMGCP:
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: galP
 C:Superfamily: glucose transport protein

[illegible]

177 CEMPETRFLITQHRROEA-TALRFMCSBQGMEDPPIGAQOSHL-----ALLR-QPGI 229
196 VLPNSRRLAEKRRIEAEVLRMLRDTSEKABEELNEIRESLKLGOGMALFKINNV 255
230 YKPFILVSIIMAFQOLSGVNAVMEFAETIPEEAKF---KDSLASVVGVLQVLETTAA 286
256 RRAVFLQMLQAMQOFPQGMNIIMYAPRIFKMAQFTTEQOMITLVYGLFPMATFIAY 315
287 LINDRAGRLLLVLSGVVWFSTSAFGAYFKLTGCGPNSSHVAISAPVSAQPDASVGL 346
316 FTVDKAGRKPAALKIGFSVMAIGTLVLG--YCLMGFDNGT-----ASSGL 357
347 AMLAVGSMCLFIAGFANGKPIPLMLSEIFPLHVKVATIGVLTNMLAFVLTKEFSS 406
358 SWLSVGMTMCIAGYASAPVWVILCSLEIOLPKCRDGFICSTTTWVSNMILGAFPLT 417
407 LMEVLRPYGAFMLASAFICFSVLTFLPCVPETKGTLEOI 446
418 LLDISGAAGTFMILTALNTAFVGTITFPLIPETKNVTLEHI 457

RESULT 9

G84864

probable membrane transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence-revision 02-Feb-2001 #text-change 16-Feb-2001
C:Accession: G84864

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.T.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.;
Neuss, D.; Nigman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, A.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487
A:Accession: G84864
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1521 <STO>
A:Cross-references: GB:AE002093; NID:g2289003; PIDN:AAB64332.1; GSPDB:GN00139
C:Genetics:
A:Gene: AT2g43330
A:Map position: 2
C:Superfamily: glucose transport protein

Query Match 24.8%; Score 577.5; DB 2; Length 521;
Best Local Similarity 31.4%; Pred. No. 3.4e-36;
Matches 148; Conservative 75; Mismatches 192; Indels 57; Gaps 9;

10 AALGPLSFGALGYSSPAISL-----ORAAPAPRLDDAASWFGAVTL---GAA 58
38 AGIGGLFEGYDTCVISCALLYIKDPEVVKOSFLOQYVNWSSFTSKLETTIVSHALVGAM 97
59 AGVILGWLVDRAGRKLSLLCSVPVAGFAVITAADVMMLGGRLITGLACGVASLVA 118
98 IGAAGGWINDYGRKKATLFAADVFAAGAIWMAADPPVILSGRLVLGVCVASSTA 157
119 PVIISLTAFAVNRLLGSCVOLMVVGLLAVLAGVVL-----EMRMLAVGCVPSLML 173
158 PVIYIAEASPEVRGCLSTVNLMTGGQFLSTLVNSAFVQVPGWKMVLGVSQPAVIOF 217
174 LMCGFHEPFRLLTQHRROEA-----ALRFLMCSBQGMEDPPIGAQOSF 219
218 ILMIFMESPRMLFMKNRKAEOVLARTYDISLDEIDHLSAE---EERKORRKTIVG 274
220 HIALLRPGIYKPIIGVSIIMAFQOLSGVNAVMEFAETIPEEAKFKDSSLASVVGVL 276
275 YLDVFRKRELRLAFACGAGIQAFOOFTGINTVMTYPTIYOMAGFHSNQLALPLSLIVA 334
277 IYVLFATAVALINDRAGRLLLVLSGVVWFSTSAFGAYFKLTGCGPNSSHVAISAPVS 336
335 MNAGATVVGIIYFIHCGRKKLAL-----SSLFGVITSL-----LLLSVSFF 375
337 AQPVDASVG--LAMLAVGSMCLFIAGFANGKPIPLMLSEIFPLHVKVATIGVLTN 394

376 KQSETSSDGLYGLWLAIVGLALYIFFAFGMPVWYNSEIIPDOYRKIGGMSATVNW 435
395 LMAFLVTEFSSLMELVRPYGAFMLASAFICFSVLTFLPCVPETKGTLEOI 446
436 ISNLIVAOFTFLIAEAGTMFFLLAGIYAVLIVFIVFVPEETGGLTFSEV 487

RESULT 10

T43400

myo-inositol transporter - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe
C:Date: 21-Jan-2000 #sequence-revision 21-Jan-2000 #text-change 20-Oct-2000
C:Accession: T43400; T39079
R:Niederberger, C.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z22484
A:Accession: T43400

A:Molecule type: DNA
A:Residues: 1575 <NIE>
A:Cross-references: EMBL:X98622; PIDN:CAA67211.1
A:Experimental source: strain 968 h90
R:Gentile, S.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21826
A:Accession: T39079
A:Molecule type: DNA
A:Residues: 290-575 <GEN>
A:Cross-references: EMBL:Z99532; PIDN:CAI61718.1; GSPDB:GN00066; SPDB:SPAC7D4.01
A:Experimental source: strain 972h-; cosmid c7D4
C:Genetics:
A:Gene: itrl1; SPDB:SPAC7D4.01
A:Map position: 1
C:Superfamily: maltose transport protein MAL61
C:Keywords: sugar transport; transmembrane protein

Query Match 24.7%; Score 574; DB 2; Length 575;
Best Local Similarity 30.5%; Pred. No. 6.9e-36;
Matches 144; Conservative 88; Mismatches 194; Indels 46; Gaps 11;

3 VFLAFAAALGPLSFGALGYSSPAISLORAAPAPRLDDAASWFGAVTLGAAGCV 62
87 VVLAFAAIGIGLFGYDTCVISCALLYIGSLVIGSL-GGHELTNGCKEFTISAFSLGALLGI 145
63 LCGWLVDRAGRKLSLLCSVPVAGFAVITAADVMMLGGRLITGLACGVASVAPYI 122
146 IAGALADFFGRKPVYIAISIIITVGSIVQVTAHHLMHMIYGRFVIGKGVGLASLIIPLYL 205
123 SIAYPAVRGLGSCVOLMVVGLLAVLAGVLE-----WRMLAVIGCVPSLMLLMC 177
206 SEIAPSKIRGRVILYVLLITAGQVIAVGDPAFBNHNMWVGLAMVPAAPQLFILI 265
178 FMPETPRFLITQHRROEA---TALRF-----LMGSQGMEDPPIGAE-----QS 218
266 WLPEPRLLVKKERQOEYNTLARTYPAHPEYIKTKLYLIDQGVADPPFSGRMKIVKT 325
219 FHLALRPGIYKPIIGVSIIMAFQOLSGVNAVMEFAETIPEEAKFKDSSLASVVGVL 278
326 FK-ELYFNPNSFRALLIACGLQAMQOLSGFNSLMFSSSTIEFVGFNNPTAAGLIIAATN 384
279 VLFATAVALINDRAGRLLLVLS--GVVWFSTSAFGAYFKLTGCGPNSSHVAISAPVS 336
385 FVFTIVAGVIDEFGRRILLTLITVGMIALIVCAVAFHF-LPKDENGNYT----- 434
337 AQPVDASVGLAMVAG--SKCLFIAGFANGKPIPLMLSEIFPLHVKVATIGVLTN 394
435 ---SGSNMAVAILVLSMTIYVAVSASGIGNLPW--QOSELFPMSVGLGISTAVNW 488
395 LMAFLVTEFSSLMELVRPYGAFMLASAFICFSVLTFLPCVPETKGTLEOI 446
489 AGNIGIGASFLTLMSEITPTGFALYGLCLFLGWLIGALFCYRDLTDITYIEE 540

Query Match	24.68;	Score 572;	DB 2;	Length 496;
Best Local Similarity	33.58;	Pred. No. 8.3e-36;		
Matches 160;	Conservative 72;	Mismatches 169;	Indels 76;	Gaps 12

```

RESULT 12      glucose transport protein 3, neuron-specific - rat
S38981
M:Alternate names: glucose transporter
C:Species: Rattus norvegicus (Norway rat)
C:Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
C:Accession: S38981
R:Nagamatsu, S.; Sawa, H.; Kamada, K.; Nakamichi, Y.; Yoshimoto, K.; Hoshino, T.
FEBS Lett. 334, 289-295, 1993
A:Title: Neuron-specific glucose transporter (NSGT): CNS distribution of GLUT3 rat glucose
A:Reference number: S38981; MUID:94063043
A:Accession: S38981
A:Molecule type: mRNA
A:Residues: 1-493 <NAG>

```

Query Match	24.5%;	Score 569;	DB 2;	Length 493;
Best Local Similarity	31.2%;	Pred. No. 1.4e-35;		
Matches 153;	Conservative 91;	Mismatches 169;	Indels 78;	Gaps 15;

RESULT 13
A11751
glucose transport protein 3 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 24-Sep-1999
C:Accession: A11751, S18128
R:Nagamatsu, S.; Kornhauser, J.M.; Burant, C.F.; Seino, S.; Mayo, K.E.; Bell, G.I.
J. Biol. Chem. 267, 467-472, 1992
A:Title: Glucose transporter expression in brain. cDNA sequence of mouse GLUT3, the b
A:Reference number: A11751; MUID:92112695
A:Accession: A11751
A:Molecule type: mRNA
A:Residues: 1-493 <NAG>
A:Cross-references: GB:61093; NID:g51088; PIDN:CAA43406.1; PID:g51089
C:Genetics:
A:Gene: GLUT3
C:Superfamily: glucose transport protein
C:Keywords: glycoprotein; transmembrane protein
#43Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	24.4%;	Score 566;	DB 2;	Length 493;
Best Local Similarity	31.3%;	Pred. No. 2.3e-35;		
Matches 153;	Conservative 86;	Mismatches 176;	Indels 74;	Gaps 13;


```

QY 216 E0SFHALLLPQPGIYKPFIIIGVSLMAFOQLSGVNAVMEFYAETIFEAKF---KDSLSAV 272
Db 259 EKIMMILCKAKTYRGLIAGVGLQVFOQFVGINTVMYSPPTIVQLAGFASNRTALLSL 318
QY 273 VVGVIQVLEFVAVALIMDRAGRRLLVLSGVVWFSTSAF-GAYFKLTGGPGNSHVA- 330
Db 319 VTAGLNAFGSIISYIFIDRIGRKKLIISLEGVIIISLGILTGVPFEATHAPAISSLETO 378
QY 331 ---ISAP-----VSAQPV-----DASY----- 344
Db 379 RFNNISCPDYKSAWNTNAMDCMTCLKASPCGYCSSPIGCKEHPGACMISDSDSVKDLCHN 438
QY 345 -----GLAMLAVGSMLFELAGFAVGCPPIPLMLSELPPLHVKGVATGICVLT 392
Db 439 ENRLMYTRGCPSPNFGFALLGLGLIIFSPGMGTVPWIVNSEIYPLRFRGICGGIATA 498
QY 393 NMLMAFLVTKFESSLMEVLRPRYGAFWLASAFCLIESVLFPLPCVPEPKGKTLLEQITAHFEG 452
Db 499 NMISNLIVAOFSLSLTAIGTSWTFELIFGVISYIALLEVMVCVPETGMPMEIEKMLER 538
QY 453 R 453
Db 559 R 559

```

Search completed: February 13, 2002, 21:53:10
 Job time: 6132 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: February 13, 2002, 21:51:44 : Search time 69.48 Seconds
(Without alignments) updates/sec
239.050 Million cell

Title: US-09-516-493-7

Perfect score: 2324
Sequence: 1 RVFVLAFAAALGPLSFGFA.....CVPETKGTLEQITAHPEGR 453

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	608.5	26.2	461	1	CSBC_BACSU
2	596	25.6	472	1	ARAE_KLEOX
3	593	25.5	464	1	GALP_ECOLI
4	589	25.3	472	1	ARAE_ECOLI
5	574	24.7	575	1	ITRL_SCHPO
6	571	24.6	494	1	GTR3_SHEEP
7	569	24.5	493	1	GTR3_RAT
8	566	24.4	493	1	GTR3_MOUSE
9	565	24.3	496	1	GTR3_HUMAN
10	560	24.1	495	1	GTR3_CANVA
11	537.5	23.1	457	1	XYLT_LACHR
12	535.5	23.0	491	1	XYLE_ECOLI
13	535	22.8	522	1	GTR2_MOUSE
14	530.5	22.8	523	1	GTR2_RAT
15	529	22.8	534	1	HDP3_CHLKE
16	529	22.8	468	1	GLCP_SYNY3
17	524	22.5	540	1	HDP2_CHLKE
18	521.5	22.4	612	1	ITR2_YEAST
19	520.5	22.4	496	1	ITR1_YEAST
20	520	22.4	522	1	ITR1_CHICK
21	518.5	22.3	584	1	ITR1_YEAST
22	517	22.2	509	1	HDP1_CHLKE
23	517	22.2	509	1	GTR4_HUMAN
24	516	22.2	509	1	GTR4_MOUSE
25	514	22.1	492	1	ITR2_MOUSE
26	512.5	22.1	523	1	ITR2_RICCO
27	511	22.0	527	1	ITR2_SCHPO
28	509	21.9	492	1	ITR1_BOVIN
29	508	21.9	492	1	ITR1_HUMAN
30	507	21.8	492	1	GTR1_RAT
31	505	21.7	490	1	GTR1_CHICK
32	504	21.7	524	1	GTR2_HUMAN
33	501	21.6	492	1	GTR1_RABIT

34	494	21.3	473	1	GLE_ZYMO
35	492.5	21.2	510	1	GTR4_MOUSE
36	489.5	21.1	533	1	GTR2_CHICK
37	485.5	20.9	510	1	HEX6_RICCO
38	484	20.8	451	1	GTR1_PIG
39	480	20.7	566	1	KHR2_KIOLA
40	475.5	20.5	482	1	YFIC_BACSU
41	473.5	20.4	547	1	GTR1_LEIDO
42	467	20.1	592	1	HXT5_YEAST
43	465	20.0	509	1	GTR4_BOVIN
44	461.5	19.9	567	1	RAG1_KIOLA
45	461.5	19.9	576	1	HXT4_YEAST

ALIGNMENTS

RESULT 1
CSBC_BACSU STANDARD; PRT; 461 AA.
ID CSBC_BACSU
AC P4633; O32289;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE METABOLITE TRANSPORT PROTEIN CSBC.
GN CSBC OR SS92BR.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=6093926; PubMed=7584049;
RA Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
RT *Cloning and sequencing of a 36 kb region of the Bacillus subtilis
RT genome between the gut and lol operons.*
RL DNA Res. 2:61-69(1995).
RN [2]
RP REVISTONS. Shibayama T., Ishio I., Aoyama D., Yoshida K.-I.;
RA Fujita Y.,
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RT Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=9930315; PubMed=10376822;
RA Akbar S., Lee S.Y., Hoylan S.A., Price C.W.;
RT *Two genes from Bacillus subtilis under the sole control of the
RT general stress transcription factor sigmaB.*
RL Microbiology 145:1069-1078(1999).
CC -I- FUNCTION: COULD SERVE EITHER A NUTRITIONAL OR AN OSMOTIC
CC PROTECTION FUNCTION.
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -I- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB005554; BAA21604.1; -
CC EMBL: Z99124; CAB16017.1; -
CC Subtilisin; BG11360; csbc.
CC InterPro: IPR003663; Sugar_transport.
CC InterPro: IPR003662; sub_transport.
CC Pfam: PF00083; sugar_tr; 1.

PRINTS: PR00171; SUGAR_TRANSPORT.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 15 35 POTENTIAL.
 FT TRANSMEM 39 59 POTENTIAL.
 FT TRANSMEM 77 97 POTENTIAL.
 FT TRANSMEM 105 125 POTENTIAL.
 FT TRANSMEM 140 160 POTENTIAL.
 FT TRANSMEM 164 184 POTENTIAL.
 FT TRANSMEM 242 262 POTENTIAL.
 FT TRANSMEM 281 301 POTENTIAL.
 FT TRANSMEM 309 329 POTENTIAL.
 FT TRANSMEM 342 362 POTENTIAL.
 FT CONFLICT 400 401 RP -> SA (IN REF. 3).
 SQ SEQUENCE 461 AA; 50235 MW; B6587B5DC7272EF1 CRC64;

Query Match 26.2%; Score 608.5; DB 1; Length 461;
 Best Local Similarity 31.8%; Pred. No. 1.2e-33;
 Matches 150; Conservative 83; Mismatches 186; Indels 53; Gaps 9;

QY 2 RVLTAAPAAALGPLSEFGALGYSSPAIPSIQRAAPAPRLDDAASWFGAVVTLGAAGC 61
 DB 6 RKMYIFFGALGLLYGYDGVISGALLFINDDIPLTITTEGLVSS---MLLGAIFGS 61
 QY 62 VLGGMVDRAGRLSLILCSVPPVAGFAVITAQDVMMLLGRLTLGLACGVAISAPV 121
 DB 62 ALSGTSDRWGRKRVVFIIFIGALACAFSGITGLIASRVILGLAVGSTALVPPY 121
 QY 122 ISEIAPAVRGLLSCVOLMVVGLAVLAGVYL---EWMVLAVLCVPSMLLMC 177
 DB 122 LSEMAPTKIRGLTGMNMLMITYGILLAVIYVLTTPPEAMWMLGLAIPAVALLIGIA 181
 QY 178 EMDTPRELTLQHRQOEAIAL-----RFLMGSEQGWEDPPIGAQSEFHLALRQ 226
 DB 182 FPEESPRLVKKSGSEEARINITHDKDIEMELAEKKQGFAP-----KETTIGVLA 236
 QY 227 PGYPRFIIIGSLMAFOOLSGVNAVMFAETIFEEAKF---KDSIASVVGVIQVLETA 284
 DB 237 KWRPMLLIGVLAIFQOAVGINTVIYAPITFKAGISASALGTMIGILNVMIC 296
 QY 285 AALIMDRAGRRLILVSGVWVWFSASGAFKLTGCGPGNSHVAISAPVAPVAVS 344
 DB 297 AMILIDVRGRKLLIMSVGTTLSLALSGVL-LTLG-----LSA 335
 QY 345 GLAWLAVGSMCLFAGAVGMPITPMLLMSTIFPLHKGVAAGICVLTNNMLAFVTK 404
 DB 336 STAMTVVPLGVIVFYQATWGPVWVWMLPELPSKARGAATGTTVLISAANLIVSL 395
 QY 405 SSMLEVLPRYGAFMLASAF---CIFSVLFTLFCVPETKGTLEOITAHFEGR 453
 DB 396 PLM---LRPMGIAMVFWFVSVICLLSEFFAFVYVPEIKGSLSEITEKSKKR 444

RESULT 2
 ARAE KLEOX
 ID ARAE KLEOX STANDARD; PRT; 472 AA.
 AC P43598;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE 15-JUL-1999 (Rel. 38, Last annotation update)
 DR ARABINOSE-PROTON SYMPORTER (ARABINOSE TRANSPORTER).
 GN ARAE.
 OS Klebsiella oxyloca.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Klebsiella.
 OX NCBI_Taxid=571;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8017;
 RX MEDLINE=95394866; PubMed=7665532;
 RA Shatwell K.P., Charalambous B.M., McDonald T.P., Henderson P.J.F.;

"Cloning, sequencing, and expression of the araE gene of Klebsiella
 oxyloca 8017, which encodes arabinose-H⁺ symport activity.";
 J. Bacteriol. 177:5379-5380(1995).
 CC - FUNCTION: UPTAKE OF ARABINOSE ACROSS THE BOUNDARY MEMBRANE WITH
 CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC - SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL; X79598; CAAS6110.1;
 DR InterPro; IPR003663; Sugar_transport.
 DR InterPro; IPR003662; sub_transport.
 DR Pfam; PF00083; sugar_tr_1.
 DR PRINTS; PR00171; SUGAR_TRANSPORT.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport.
 FT TRANSMEM 23 43 POTENTIAL.
 FT TRANSMEM 63 83 POTENTIAL.
 FT TRANSMEM 92 110 POTENTIAL.
 FT TRANSMEM 120 138 POTENTIAL.
 FT TRANSMEM 149 167 POTENTIAL.
 FT TRANSMEM 177 197 POTENTIAL.
 FT TRANSMEM 259 278 POTENTIAL.
 FT TRANSMEM 299 318 POTENTIAL.
 FT TRANSMEM 328 348 POTENTIAL.
 FT TRANSMEM 362 382 POTENTIAL.
 FT TRANSMEM 396 415 POTENTIAL.
 FT TRANSMEM 424 443 POTENTIAL.
 SQ SEQUENCE 472 AA; 51732 MW; 410021E1BEE3D96E CRC64;

Query Match 25.6%; Score 596; DB 1; Length 472;
 Best Local Similarity 33.7%; Pred. No. 8.3e-33;
 Matches 153; Conservative 84; Mismatches 181; Indels 36; Gaps 9;

QY 7 AFPAALGPLSEFGALGYSSPAIPSIQRAAPAPRLDDAASWFGAVVTLGAAGVLCGW 66
 DB 26 STAAAVAGLLFGDLGDIAGALPFTIDHFLVSSRLQ---WVSSMMLGAALGALFNGW 81
 QY 67 LVDRAGRLSLILCSVPPVAGFAVITAQDVMMLLGRLTLGLACGVAISAPVYIS 126
 DB 82 LSFRLGRKRSMLWGVAVLFVAGSVGSAPATSVMLLVARIIVGVAAGIASYAPVLSMA 141
 QY 127 YPAVGILGSCVOLMVVGLIAYLAGVLEW---RWLAVLG---CVPSMLLMCMPE 182
 DB 142 SENVRGKMISMYQLMVTGIIYMAFLSDPAFSYSGWBRMLGVLPAVYLIITVPLNS 201
 QY 183 PFFLTQHRQEA-IALPFLMGSEQGWEDPPIGAQSEHL-----ALR-OPGIYKPT 235
 DB 202 PRLIAEKGRHVEAEVFLMLDITSEKARDELNETRESIKLKQGGMALEKVRNRRVAV 261
 QY 236 CVSLMAFOOLSGVNAVMFAETIFEEAKF---KDSIASVVGVIQVLETAVALIM 292
 DB 262 GMLLOAMQOFTGMNIMYAPRIKMGFTTTEGQWATLVGLTFPFAIFIAVFDKA 321
 QY 293 GRRLIVLSGVVWVWFSASGAFKLTGCGPGNSHVAISAPVAPVAVSAPVAVSGL 354
 DB 322 GRKALVIGSVMAIGTLVGL---YCLMQDFNGT-----ASSGLSMLSV 354
 QY 353 SMCFLIAGFVAGMPITPMLLMSTIFPLHKGVAAGICVLTNNMLAFVTKES 354
 DB 364 MTMNCIAGVAMSAPVWMLCSEIOLPKCDPGITSTTNWVSNMIIGATP 354
 QY 413 PYGAFWLASFCIFSVLFTLFCVPETKGTLEOI 446

```

Db      424 AACFMVLYTALNVAFLGVTEFWLIPETKNVLEHT 457
RESULT  3
ID      GALP_ECOLI          STANDARD:          PRT:          464 AA.
AC      P37021;
DT      01-JUN-1994 (Rel. 29, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DT      20-AUG-2001 (Rel. 40, Last annotation update)
DE      GALACTOSE-PROTON SYMPORTER (GALACTOSE TRANSPORTER).
GN      GALP OR B2943.
OS      Escherichia coli.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
OX      NCBI_TaxID=562;
RN      [1]
RP      SEQUENCE FROM N.A., AND CHARACTERIZATION.
RL      Roberts P.E.;
RT      Thesis (1992), University of Cambridge, U.K.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K12 / MG1655;
RX      MEDLINE=97426617; PubMed=9278503;
RA      Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Shao Y.;
RT      "The complete genome sequence of Escherichia coli K-12.";
RL      Science 277:1453-1474(1997).
CC      -1- FUNCTION: UPTAKE OF GALACTOSE ACROSS THE BOUNDARY MEMBRANE WITH
CC      THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC      -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
CC      -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb.sib.ch).
CC      -----
DR      EMBL: 028377; AAA69110.1; -
DR      EMBL: AE000377; AAC75980.1; -
DR      EcoGene: EG12148; galp.
DR      InterPro: IPR003663; Sugar_transporter.
DR      InterPro: IPR00083; sub_transporter.
DR      Pfam: PF00083; sugar_tr.1.
DR      PRINTS: PR00171; SUGRTNSPORT.
DR      PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR      PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW      Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
KW      Complete proteome.
FT      TRANSMEM 16 36
FT      TRANSMEM 57 77
FT      TRANSMEM 85 105
FT      TRANSMEM 113 133
FT      TRANSMEM 140 160
FT      TRANSMEM 172 192
FT      TRANSMEM 251 271
FT      TRANSMEM 291 311
FT      TRANSMEM 322 342
FT      TRANSMEM 352 372
FT      TRANSMEM 395 415
FT      TRANSMEM 417 437
FT      TRANSMEM 464 AA; 50982 MW; 07E08935BD8E3F8E CRC64;
SQ      SEQUENCE

```

```

Query Match      25.5%; Score 593; DB 1; Length 464;
Best Local Similarity 32.6%; Pred. NO. 1,3e-32;
Matches 156; Conservative 83; Mismatches 173; Indels 66; Caps 12;

```

```

QY      4 FLAFAAALGPLSFQFALGYSSPAIPSLQRAAPAPRDLDDAAASWFGAVVTLGAAGCVL 63
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      16 FFFVFLAALAGLLFGDLGVIAGALPFIADFF---QITSHQEWVSSMMGGAAGVAGV 71
QY      64 GCVLVDRAGRKLSLLSCVPPVAGFAVTTAADVMMLLGGRLLTGLAGVASLAPVYIS 123
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      72 SGWLSPFKLGRKRSIMIGAILTFVAGSLFSAAPANEVLLSLVGLAGVASVYAPLPLS 131
QY      124 EIAVPAVAGLLSGCVQLMWWVVGILAYLA----GVLMEKRLVAGVPPSILMLMGFM 179
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      132 EIAPEKTGSMISMVQMLMTITGILGAVLSDTAFSTGAMRMNLGVIIIPAILLLIGVFPL 191
QY      180 PETPPELLTQHR--RQEAIALRF-----LWSEQGWEDDPDGAEGSFH 220
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      192 PDSPPMFPAKRFRVDAERVLRLRDTSAEKRELDREISLQVKSQW----- 239
QY      221 LALLRQGITVPPFI-IGSLMAFOOLSGVNMVMPFAETIFPEAKRKDS---LASVYGV 276
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      240 -ALFRENSEFRRAVFLGLVGLVDMODFTGMNVIYAPKIPELAGVTNTTEOMMGTVIGL 298
QY      277 IQVLTAAVALIMPRAGRLLVLGVVNVFSTSAFGAVFKLTQGGPNSHVAISAPVS 336
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      299 TNNVLTATFALIGLVDKMGKPLTLTGFLVM-----AAG-----KGVLTGMHIGIHSP-S 346
QY      337 AQPVDASVGLAMLAAGSNCLEFIAGFVAGVGPIPWLMLSELPILHYKVATGICVLTNNMLM 396
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      347 AQ-----YFALAMLMFIVGFAMSAGPLIWLGCSEIPLDKGRDPCITCSTATNMIA 397
QY      397 AFLVTKERSSLMFVLRPGAFWLASFCIFSVLPLFCVPEPKKGTLEQITAH-PEGR 453
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      398 NMIVGAFELTMLNLTGNANTFVYAAALNVLEPLTLMLVPEPKHVSLEHIERNLMKGR 455
RESULT  4
ID      ARAE_ECOLI          STANDARD:          PRT:          472 AA.
AC      P09830; Q46937;
DT      01-MAR-1989 (Rel. 10, Created)
DT      01-MAR-1989 (Rel. 10, Last sequence update)
DT      20-AUG-2001 (Rel. 40, Last annotation update)
DE      ARABINOSE-PROTON SYMPORTER (ARABINOSE TRANSPORTER).
GN      ARAE OR B2841 OR Z4161 OR ECS3698.
OS      Escherichia coli, and
OS      Escherichia coli 0157:H7.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
OX      NCBI_TaxID=562, 83334;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=87115869; PubMed=3543693;
RA      Maiden M.C.J., Davis E.O., Baldwin S.A., Moore D.C.M.,
RA      Henderson P.J.F.;
RT      "Mammalian and bacterial sugar transport proteins are homologous.";
RT      Nature 325:641-643(1987).
RL      Nature 325:641-643(1987).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K12 / JM2433;
RX      MEDLINE=88228015; PubMed=2836407;
RA      Maiden M.C.J., Jones-Mortimer M.C., Henderson P.J.F.;
RT      "The cloning, DNA sequence, and overexpression of the gene araB
RT      coding for arabinose-proton symport in Escherichia coli K12.";
RL      J. Biol. Chem. 263:8003-8010(1988).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K12 / MG1655;
RX      MEDLINE=97426617; PubMed=9278503;
RA      Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Shao Y.;
RT      "The complete genome sequence of Escherichia coli K-12.";
RL      Science 277:1453-1474(1997).
RN      [4]

```

RC SEQUENCE FROM N.A.
 RC STRAIN-0157:H7 / EDL933 / ATCC 700927;
 RA MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postel G., Hackett J., Klink S., Boulton A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocistus K.,
 RA Apodaca J., Mantharman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7,"
 RL Nature 409:529-533(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-0157:H7 / RMD 0509952;
 RA MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino E., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Ikeda T., Takami H., Honda T., Sasaki C., Ogasawara T.,
 RA Kubara S., Shiba T., Hattori M., Shinagawa H.,
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12,"
 RL DNA Res. 8:11-22(2001).
 RN [6]
 RP PRELIMINARY SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE=84114868; PubMed=6319708;
 RA Stoner C., Schleif R.F.;
 RT "The area low affinity L-arabinose transport promoter. Cloning,
 RT sequence, transcription start site and DNA binding sites of
 RT regulatory proteins,"
 RL J. Mol. Biol. 171:369-381(1983).
 CC -1- FUNCTION: UPTAKE OF ARABINOSE ACROSS THE BOUNDARY MEMBRANE WITH
 CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J03732; AAA23469.1; -
 DR EMBL: X00272; CA25075.1; -
 DR EMBL: U29581; AAB4048.1; -
 DR EMBL: AE000368; AAC75880.1; -
 DR EMBL: AE005513; AAC57953.1; -
 DR EMBL: AP002563; BAB37121.1; -
 DR PIR: B26430; B26430.
 DR PIR: A28075; A28075.
 DR Ecogene; EG10056; araE.
 DR InterPro: IPR003663; Sugar_transp.
 DR InterPro: IPR003662; sub.transp.
 DR Pfam; PF00083; sugar_tf.1.
 DR PRINTS; PR00171; SUGRTNSPRT.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
 KW Complete proteome.
 FT TRANSMEM 23 43 POTENTIAL.
 FT TRANSMEM 63 83 POTENTIAL.
 FT TRANSMEM 92 110 POTENTIAL.
 FT TRANSMEM 120 138 POTENTIAL.
 FT TRANSMEM 149 167 POTENTIAL.
 FT TRANSMEM 177 197 POTENTIAL.
 FT TRANSMEM 259 278 POTENTIAL.
 FT TRANSMEM 299 318 POTENTIAL.
 FT TRANSMEM 328 348 POTENTIAL.
 FT TRANSMEM 362 382 POTENTIAL.
 FT TRANSMEM 396 415 POTENTIAL.
 FT TRANSMEM 424 443 POTENTIAL.
 FT CONFLICT 26 SVA -> YDR (IN REF. 6).

SO SEQUENCE 472 AA; 51684 MW; 411990A441D44393 CRC64;
 Query Match 25.38; Score 589; DB 1; Length 472;
 Best Local Similarity 33.58; Pred. No. 2,4e-32;
 Matches 154; Conservative 81; Mismatches 189; Indels 36; Gaps 9;
 OY 1 RRVLAFAAALGLPSGFALGYSSPAIPSLQRAAPPAPRLDDAASWGAAYTLCGAAG 60
 DB 20 RMNFVSAVAAGVGLFDLDIGVYAGALPTTTHFVLTSLQD---VWSSMGLCAAIG 75
 OY 61 GVLGMLVDPACKRLSLILCSVPFVAGFAVTAADVMMLGRLTLGLACVAAIAAVY 120
 DB 76 ALNNGWISFRLGRKYSIMAGAILFVLGSGAFATSEVMLIARVYLAVIGASTAPL 135
 OY 121 YISEIAYPAVRGLGSCVQLMVVGLLAVIAGVLEW--RMLAVLG--CVPSLMILL 176
 DB 136 YLSEMASENVRGIMIMYQMLMTGLVLAFLSDTAFSYSGNRRAMLGVLAFLVILLY 195
 OY 177 CFMPETPRLLTOHROEA-IALRFLMGSEQMEDPPIGAESFHL-----ALLR-QPGI 229
 DB 196 VFLPNSPRLIAKRGRIEAEVILRMLRDYSEKAREELNRESLTKOGGMALEKIRNV 255
 OY 230 YPFILIGVSLAPFOOLSGVNAVMEFAETFEFAK---KDSLAVVGVITQVTAFAA 286
 DB 256 RRAVFLGMLLQAMQOFTGNIIITVYAPRIEFKAGFTTEQOIMATLVGLTFMEFTFAV 315
 OY 287 LMDRAGRLLLVLSGVVWFSTSAFAYFKLTQGGPNSSHVATSAVSAQPDASVGL 346
 DB 316 FTVDPAGKRPALKIGFVSVALGLTVLG--YCLMQPDNT-----ASSGL 357
 OY 347 AMLAAGSMCLFAGFAVAGMGPILMLSEIFPLHVKVATGICVLTMLMAFLTKRESS 406
 DB 358 SMLSTGVMCMCIAGYAMSAAPVWMLCSEIQLPCRFGITCTTFTWVSMIIGAFELT 417
 OY 407 LMEVLRPGAFWLASAFICFSVFLTFCVPEPKGKTLEQI 446
 DB 418 LLDSIGAAGTFLYALNIAVGLTFMLIPETKAVTLEHI 457
 RESULT 5
 ID ITRL_SCHPO STANDARD; PRT; 575 AA.
 AC Q10286;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE MYO-INOSITOL TRANSPORTER 1.
 GN ITR1 OR SPACT7D4.01.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=968_H90;
 RX MEDLINE=98228265; PubMed=9560432;
 RA Niederberger C., Grab R., Schweingruber A.-M., Fankhauser H.,
 RA Rusu M., Politea M., Edenharter L., Schweingruber M.E.;
 RT "Exogenous inositol and genes responsible for inositol transport are
 RT required for mating and sporulation in Schizosaccharomyces pombe,"
 RL Curr. genet. 33:255-261(1998).
 RN [2]
 RP SEQUENCE OF 290-575 FROM N.A.
 RC STRAIN=972;
 RA Gentles S., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;
 RT Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RL -1- FUNCTION: TRANSPORTER FOR MYO-INOSITOL.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

FT CARBOHYD 43 43 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 494 AA: 54194 MW: A89204D3EA74BBBA CRC64:

Query Match 24.6%; Score 571; DB 1; Length 494;
Best Local Similarity 30.5%; Pred. No. 4e-31;
Matches 149; Conservative 94; Mismatches 179; Indels 66; Gaps 10;

OY 3 VFLLAFAALGPIISFGFALGYSSPAIRSL-----ORAPPAPRLDDAAASWFGAV 52
DB 10 LIFAISATIGSFQFGNTVINAPPAIIPDFINLYLEESSETPSSVLSMSVAI 69
OY 53 VITGAAGVILGGMVDRAGKSLILCSVPYAGFAVI---TAADPVMMLGGRLTGL 109
DB 70 FSVGGMIGSFVGLFVNRFRNSMLIVNLALAGGLMGFCMAEVEVLLIGRLTGL 129
OY 110 AGCVASIVAPVYISEIAYPAVRGLGSCVOLMVVGYLLAYLAGVWL----EMRWLAVLG 165
DB 130 FCGIGTGFVPMYIGETISPTALRGAFCTLNOLGIVIGILVAQIFGLKVLGTEIDIMPLL 189
OY 166 --CVPESLMLLMCFMPEPRFLITTOHRROEAT--ALRFIMGSGQWEDPPIGAOSFH- 220
DB 190 FTIIPALIGCAALPECPSPFLINRKEEKAKELLQRLMGTEVDVADIOEMKDESRM 249
OY 221 -----LALLRPGIYKPFILIGVSIMAFQOLSGVNAVMFAETTFEAKFKDSLASV 272
DB 250 SOEKQVIVLELFRAPNRQPIIISIMOLSOQLSGINAVYYSIGIFKDAVGQEPVAT 309
OY 273 VGVYIOVLFATAVAALIMDRAGRLLLVLSGVVWVSTSAFGANFKLTGGCGNSSHVAIS 332
DB 310 GAGVNTLFTVVSFVLEVRAGRRLHLI-----GIGGMAFCISIL 348
OY 333 APVSAQVVDASVGLAMLVAGSMCLFTGPAVGMGPIMLMSEIFLHVGVATGICVLT 392
DB 349 MTISLLIKDNYSWMSFCIGILVFAFEIGPPIWFIYAELEFGCGPRAMAANAAGCS 408
OY 393 NMLAPLVTKFEFSSIMELVRPYGAFWL--ASAFCLFSV-----LFTLVCVPEKGTLEQ 445
DB 409 NMTSNFLV-----GILPSPATFYLCAVYFIVETVLFVWFTEFKVETRGRTFEE 460
OY 446 ITAHFEGR 453
DB 461 ITRAFFEGO 468

RESULT 7
GTR3_RAT
ID GTR3_RAT STANDARD: PRT: 493 AA.
AC 007647: 062729:
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-1998 (Rel. 36, Last annotation update)
DE GLUCOSE TRANSPORTER TYPE 3, BRAIN.
GN SLC2A3 OR GLUT3 OR GLUT-3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94063043; PubMed=8243635;
RA Nagamatsu S., Sawa H., Kamada K., Nakamichi Y., Yoshimoto K.,
RA Hoshino T.;
RT "Neuron-specific glucose transporter (NSGT): CNS distribution of
RT GLUT3 rat glucose transporter (RGF3) in rat central neurons";
RL FEBS Lett. 334:289-295(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=96042338; PubMed=7475896;
RA Krishnan S.N., Haddad G.G.;
RT "Cloning of glucose transporter-3 (GLUT3) cDNA from rat brain";
RL Life Sci. 56:1193-1197(1995).

CC - FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. PROBABLY A NEURONAL
CC GLUCOSE TRANSPORTER.
CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC - TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN.
CC - SIMILARITY: HIGH TO OTHER GLUCOSE TRANSPORTERS.
CC - SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@isb-sdb.ch).
DR EMBL: D13962; BAA03065.1; -
DR PIR: J17978; AAB62503.1; -
DR PIR: S38981; S38981.
DR InterPro: IPR003663; Sugar_transprtr.
DR InterPro: IPR003662; sub_transprtr.
DR Pfam: PF00083; sugar_tr.1.
DR PRINTS: PR00171; SUGRTRNSPORT.
DR PRINTS: PR00172; GLUCTRNSPORT.
DR PRINTS: PR01192; GLUCTRNSPORT3.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transporter; Transport; Glycoprotein;
KW Multigene family.
FT TRANSMEM 11 31 POTENTIAL.
FT DOMAIN 32 64 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 65 85 POTENTIAL.
FT TRANSMEM 94 114 POTENTIAL.
FT TRANSMEM 125 145 POTENTIAL.
FT TRANSMEM 154 174 POTENTIAL.
FT TRANSMEM 184 204 POTENTIAL.
FT DOMAIN 205 269 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 270 290 POTENTIAL.
FT TRANSMEM 305 325 POTENTIAL.
FT TRANSMEM 336 355 POTENTIAL.
FT TRANSMEM 367 387 POTENTIAL.
FT TRANSMEM 400 420 POTENTIAL.
FT TRANSMEM 428 448 POTENTIAL.
FT DOMAIN 449 493 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 493 493 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 447 447 R -> S (IN REF. 2).
FT CONFLICT 447 447 S -> F (IN REF. 2).
FT CONFLICT 470 470 H -> Q (IN REF. 2).
SQ SEQUENCE 493 AA: 53580 MW: 745446B59BDF4399 CRC64:

Query Match 24.5%; Score 569; DB 1; Length 493;
Best Local Similarity 31.2%; Pred. No. 5.4e-31;
Matches 153; Conservative 91; Mismatches 169; Indels 78; Gaps 15;

OY 6 AAFPAALGPIISFGFALGYSSPAIRSLORAPP-----PRDDDAASWFGAVY 54
DB 13 AVFVATIGSFQFGNTVIN-----APETIKDFNLYLEERLEDPRE--GLTTP 61
OY 55 LGA-----AAGVILG-----GMLVDRAGKSLILCSVPV-----AGFAVITAADVMM 99
DB 62 LWSICVAIFSYGGMIGSFVGLFVNRFRNSMLIVNLALAGGLMGFPAKI--AESEVEM 119
OY 100 LLAGRLTLTGACVAVIAPVYISEIAYPAVRGLGSCVOLMVVGYLLAYLAGVWL--WVL- 156
DB 120 LILGRLLIGIFCGIGTGFVPMYIGESPTALRGAFCTLNOLGIVIGILVAQVGLDPITLG 179
OY 157 -EMRWLAVLG--CVPESLMLLMCFMPEPRFLITTOHRROEATA--LRFIMGSE----- 205
DB 180 SEELMPILGLITIPALIGCAALPECPSPFLINRKEEKAKELLQRLMGTEVDVADIOEI 239
OY 206 OGWEDPPIGAOSFH--LALLRPGIYKPFILIGVSIMAFQOLSGVNAVMFAETTFEEA 262
DB 240 OEMKDESRMSQEKQVIVLELFRKSPSYFQPIIISVIVLQSOQSGINAVYYSIGIFGQDA 299

AC P11169:
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE GLUCOSE TRANSPORTER TYPE 3, BRAIN.
 GN SLC2A3 OR GLUT3.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=69008414; PubMed=3170580;
 RA Kayano T., Fukumoto H., Eddy R.L., Fan Y.-S., Byers M.G., Shows T.B.,
 RA Bell G.L.;
 RT "Evidence for a family of human glucose transporter-like proteins.
 RT Skeletal muscle and other tissues."
 RL J. Biol. Chem. 263:15245-15248(1988).
 CC - FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. PROBABLY A NEURONAL
 CC - GLUCOSE TRANSPORTER.
 CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC - TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN.
 CC - SIMILARITY: HIGH TO OTHER GLUCOSE TRANSPORTERS.
 CC - SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: M20681; AAB61083.1; -
 DR PIR: A31986; A31986.
 DR MIM: 138170; -
 DR InterPro: IPR003663; Sugar_transporter.
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr.1
 DR PRINTS: PR00171; SUGTRNSPORT.
 DR PRINTS: PR00172; GLUCTRNSPORT.
 DR PRINTS: PR01192; GLUCTRNSPORT3.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane; Sugar transporter; Transport; Glycoprotein;
 KW Multigene family.
 FT DOMAIN 11 31 POTENTIAL.
 FT TRANSMEM 32 64 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 65 85 POTENTIAL.
 FT TRANSMEM 94 114 POTENTIAL.
 FT TRANSMEM 125 145 POTENTIAL.
 FT TRANSMEM 154 174 POTENTIAL.
 FT TRANSMEM 184 204 POTENTIAL.
 FT DOMAIN 205 269 POTENTIAL.
 FT TRANSMEM 270 290 POTENTIAL.
 FT TRANSMEM 305 325 POTENTIAL.
 FT TRANSMEM 336 355 POTENTIAL.
 FT TRANSMEM 367 387 POTENTIAL.
 FT TRANSMEM 400 420 POTENTIAL.
 FT TRANSMEM 428 448 POTENTIAL.
 FT DOMAIN 449 496 POTENTIAL.
 FT CARBOHYD 43 43 CYTOPLASMIC (POTENTIAL).
 FT SEQUENCE 496 AA: 53924 MW: F601C6D892FE16516 CRC64;
 .SQ

Query Match 24.3% Score 565; DB 1: Length 496;
 Best Local Similarity 31.4% Pred No. 9.9e-31;
 Matches 151; Conservative 88; Mismatches 190; Indels 52; Gaps 9;

OY 3 VFLAARAAALGPLESGFALGYSSPAIPSLOR-----AAPAPRLDDAAASWFGAV 52
 DB 10 LIFALIVATIGSFQFGYNGVINAEKIKERLTKTLTKDGNAPSEVLLTSLMSLSVAI 69

OY 53 VTLGAAGVYLGLVDRAGRKSLICSPFVAG---FAVITAADVWMLGRLITGL 109
 DB 70 FSVGMATIGSSVGLFVNRFRGRNSMLVLNLAATGCGFMGLCKVAKSVKLLIGRLVIGL 129
 OY 110 ACQVASLVAPEVYISETAPPAVRGILGSCVQDMVAVVIGILAYLAG--WVL--EMRLAVLG 165
 DB 130 FCGICIGFVPMYIGELISPTALRCAPFGLNGLGIVGILVAQITGLEITIGSEELPGLLG 189
 OY 166 --CVPSIMALLMCFMPEIRPRLTLTHRQEA--ALRELMGSEQGMEDPPIGAEGSFH- 220
 DB 190 FTLLPAILGSAALPFCPESEPRFLINRKEENAKOILLRLMGVQDSQIQMKDESARM 249
 OY 221 -----LALLRPGIYKPEITIGVSLMFPQOLSGVNAWMAETLFEAKRKDSLAV 272
 DB 250 SQEKQVTVLELFVSSYRPIIISYLOLSQOLSGVNAWMAETLFEAKRKDSLAV 309
 OY 273 VGVGVTVLETAVALIMDRAGRLLLVSGVWVSTSAFAYFKLTGGPNSHVAIS 332
 DB 310 GAGVWNTITFVSLFLVERAGRRLHMI-----GLGMAFCSTL 348
 OY 333 APVSAQPVDAVSGIAMIAVSMCLFIAGFVAGWGPIMVLSLPIPLHVKVATGICVLT 392
 DB 349 MIVSLILNDNTNGMFCVIGALIVFAFEIGPGIPIWFIYAEUFSQGRPDAAMAAGCS 408
 OY 393 NMLMAFLVTKERSSLMELVRPGAFWLASAFICISVLTFLCPVETKGTLEOTTAHEG 452
 DB 409 NMTSMELVGLLFPSSAHYLGAV-VFLIFTGFLITFLATTFKVPETKRTFEDITRAEG 467
 OY 453 R 453
 DB 468 Q 468

RESULT 10
 GTR3_CANFA
 ID GTR3_CANFA STANDARD; PRT; 495 AA.
 AC P47842;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE GLUCOSE TRANSPORTER TYPE 3, BRAIN.
 GN SLC2A3 OR GLUT3.
 OS Canis familiaris (Dog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN 11
 RP SEQUENCE FROM N.A.
 RX TISSUE=Brain cortex;
 RX MEDLINE=96194907; PubMed=8654954;
 RX Borson N.D., Salo W.L., Drewes L.R.;
 RT "Canine brain glucose transporter 3: gene sequence, phylogenetic
 RT comparisons and analysis of functional sites."
 RL Gene 168:251-256(1996).
 CC - FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. PROBABLY A NEURONAL
 CC - GLUCOSE TRANSPORTER.
 CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC - TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN.
 CC - SIMILARITY: HIGH TO OTHER GLUCOSE TRANSPORTERS.
 CC - SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: L35267; AAA51454.1; -
 DR InterPro: IPR003663; Sugar_transporter.
 DR InterPro: IPR003662; sub_transporter.


```

Db 294 VTAIAVALMDKIDRKKIYNIGAVGMSISLFVMSIGMKFS-----GSSQNAI----- 340
QY 341 DASVCIAMLAAGSMCLFAGFVAGCGPIPMILMSIFPLAHKVAATGICVLTNMIMAFV 400
Db 341 -----ISVIALTVIAFFSATWGPVWVWVIGSEVPLNRGLGNFSVYINWNTAMIV 392
QY 401 TKRESSIMEVLPRYGAFWMLASAFICFVSFLFTFCVPEPKCKTLPOITA 448
Db 393 SLTFPSLIDFPGTSLFTGVCILCEASIMFWOKRVFETFRNSLSDIRA 440

RESULT 12
XYLE_ECOLI
ID XYLE_ECOLI STANDARD; PRT; 491 AA.
AC P09098;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE D-XYLOSE-PROTON SYMPORTER (D-XYLOSE TRANSPORTER).
GN XYLE OR B4031 OR Z5629 OR ECS5014.
OS Escherichia coli, and
OC Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA MEDLINE=88007632; PubMed=2820984;
RA Davis E.O., Henderson P.J.F.;
RT "The cloning and DNA sequence of the gene xyle for xylose-proton
RT symport in Escherichia coli K12."
RL J. Biol. Chem. 262:13928-13932(1987).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=87115869; PubMed=3543693;
RA Marden M.C.J., Davis E.O., Baldwin S.A., Moore D.C.M.,
RA Henderson P.J.F.;
RT "Mammalian and bacterial sugar transport proteins are homologous."
RL Nature 325:641-643(1987).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=94089392; PubMed=8265357;
RA Blatter F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT region from 89.2 to 92.8 minutes."
RL Nucleic Acids Res. 21:5408-5417(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimantla E.T., Potamoudis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blatter F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kusata S., Shiba T., Hattori M., Shimagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).

```

```

RN [6]
RP SEQUENCE OF 1-192 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=88234001; PubMed=2836810;
RA Francoz E., Dassa E.;
RT "3' end of the malefG operon in E.coli: localization of the
RT transcription termination site."
RL Nucleic Acids Res. 16:4097-4109(1988).
CC -1- FUNCTION: UPTAKE OF D-XYLOSE ACROSS THE BOUNDARY MEMBRANE WITH
CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
CC -1- INDUCTION: BY XYLOSE.
CC -1- MISCELLANEOUS: E.COLI HAS TWO D-XYLOSE TRANSPORT SYSTEMS THAT
CC ACCUMULATE SUGAR AGAINST A CONCENTRATION GRADIENT: THE XYLE SYSTEM
CC WHICH UTILIZES THE ELECTROCHEMICAL GRADIENT OF PROTONS AND THAT IS
CC INSENSITIVE TO COLD OSMOTIC SHOCK AND THE XYLE SYSTEM THAT USES A
CC HIGH-ENERGY PHOSPHATE COMPOUND AND IS SENSITIVE TO COLD OSMOTIC
CC SHOCK.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J02812; AAA9016.1; -
DR EMBL: U00006; AAC43125.1; -
DR EMBL: AE000476; AAC77001.1; -
DR EMBL: AE005636; AAC59230.1; -
DR EMBL: AP002568; BAB38437.1; -
DR EMBL: X06663; CAA29863.1; -
DR PIR: A26430; A26430.
DR PIR: A27418; A27418.
DR PIR: S00874; S00874.
DR EcoGene: EG11076; XYLE.
DR InterPro: IPR003663; Sugar_transport.
DR InterPro: IPR003662; sub_transport.
DR Pfam: PF00063; sugar_tr; 1.
DR PRINTS: PR00171; SUGRTNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
DR KX Complete proteome.
KW Complete proteome.
FT FT TRANSMEM 12 32
FT FT TRANSMEM 57 77
FT FT TRANSMEM 91 111
FT FT TRANSMEM 135 155
FT FT TRANSMEM 167 187
FT FT TRANSMEM 202 222
FT FT TRANSMEM 274 294
FT FT TRANSMEM 314 334
FT FT TRANSMEM 344 364
FT FT TRANSMEM 371 391
FT FT TRANSMEM 409 429
FT FT TRANSMEM 445 465
FT FT CONFLICT 64 64
SQ SQUENCE 491 AA; 53608 MW; 2AFLAF9756C0B722 CRC64;

```

Query Match 23.0%; Score 535.5; DB 1; Length 491;

Best local Similarity 30.1%; Pred. No. 8.9e-29; Matches 151; Conservative 87; Mismatches 180; Indels 83; Gaps 15;

```

QY 3 VFPLAAFAALGLPLSFALGYSPPAIPSIQRAAPAPRLDDAAA-SWEG---AVYTLGAA 58
Db 10 IFSITLVALTGLGLFSDYAVISGVESLNTVFVAQPNSESAAANSLLCFVAASALICGI 69
QY 59 AGGVLCGMVLVDAGRKLSLLCSVPFV-----AGFAVITAAQDVMMLLG----- 103
Db 70 TCGALGGYCSNNFGRRDSIKIAVAVLFITSGVSAWPELGFSTINPDNTVPYTLAGVPEF 129

```

```

OY 104 ---RLTLGLACVASLAPVYISFIAPPAVAGLIGSCVOLMVVIGILLAYL----- 151
DB 130 VVIRIIGIGVGLASMLSPWYIAELAPAHIRGKIVSFNOFAIFGOLLVCVNYFIARSG 189
OY 152 -ACWVLE--WRMLAVICGVPSRLMLLMCEMPETPRFLLOHROEALIA-LRFLMGSPOG 207
DB 190 DASWLTNDGKRIRFASCEIFALLFLMLLYVPSPPWMLMSKGOEALGKLTMGNTLA 249
OY 208 MEDPPIGAESFHAL-----LRQPGIYKPFILIGVSLMAFOOLSGVNAVFYAETI 258
DB 250 TQ-----AVGEIKHSLDHGKRKGGRLMPCV-GVIVIGWMLSTFQGVGIVNVLVYAEV 303
OY 259 FEE--AKFQDSSLASVYVGVYQVLFVAVALMDRAGRLLLVLSGVVWFSTSAFQAYF 316
DB 304 FKTLGASTDIALLOTIIIVGVINLFTVLAIMTVDKFGKRPLQIIGALGMALGMSFLGTA 363
OY 317 KLTOGGGNSHVAISAPVSAQVVDASVGLAMLAVGSMCFIFIGFVAGWGPIMPILMSEI 376
DB 364 -YTO-APG-----IYALLSMLEFVAAFAMSWGCVWVLLSEI 398
OY 377 FPLHVGAVATGICVLTNMLMAFLVTKEF-----SSLMEVLRPYGAFMLASFCIFSVLF 430
DB 399 FPMALIGKALAIIVAAQWOLANVFSWTFPRMMDKNSMLVAHFHNGFSYVITGCKGVLAALF 458
OY 431 TLFCVPETKGKTLQDTAHEF 451
DB 459 MKFVPETKGTLELEALWE 479

```

RESULT 13

```

GTR2_RAT 13
ID GTR2_RAT STANDARD: PRT: 522 AA.
AC P12336:
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DN GLUCOSE TRANSPORTER TYPE 2, LIVER.
OS SLC2A2 OR GLUT-2.
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=89003066; PubMed=3048704;
RA Thorens B., Sarkar H.K., Kaback H.R., Lodish H.F.;
RT Cloning and functional expression in bacteria of a novel glucose
RT transporter present in liver, intestine, kidney, and beta-pancreatic
RT islet cells.
RL Cell 55:281-290(1988).
CC - FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY
CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
CC MEMBRANE OF HEPATOCYTES & IS RESPONSIBLE FOR UPTAKE OF GLUCOSE BY
CC THE BETA CELLS. IT MAY COMPRISE PART OF THE GLUCOSE-SENSING
CC MECHANISM OF THE BETA CELL. IT MAY ALSO PARTICIPATE WITH THE
CC NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSELLULAR TRANSPORT OF
CC GLUCOSE IN THE SMALL INTESTINE & KIDNEY.
CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE; PROPEIN.
CC - TISSUE SPECIFICITY: PRESENT IN LIVER, INTESTINE, KIDNEY AND
CC BETA-PANCREATIC ISLET CELLS.
CC - SIMILARITY: HIGH TO OTHER GLUCOSE TRANSPORTERS.
CC - SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC - This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: J03145; AAA1298.1; -

```

```

DR PIR: A31556.
DR InterPro: IPR003663; Sugar_transprtr.
DR InterPro: IPR003662; sub_transprtr.
DR Pfam: PF00083; sugar_tr_1.
DR PRINTS: PR00171; SUGTRNSPORT.
DR PRINTS: PR00172; GLUCTRNSPORT.
DR PRINTS: PR01191; GLUCTRNSPORT2.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Transport; Glycoprotein;
KW Multigene family.
FT TRANSMEM 11 31 POTENTIAL.
FT DOMAIN 32 97 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 98 117 POTENTIAL.
FT TRANSMEM 126 146 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 186 206 POTENTIAL.
FT TRANSMEM 216 236 POTENTIAL.
FT DOMAIN 237 301 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 302 322 POTENTIAL.
FT TRANSMEM 337 357 POTENTIAL.
FT TRANSMEM 368 388 POTENTIAL.
FT TRANSMEM 399 419 POTENTIAL.
FT TRANSMEM 432 452 POTENTIAL.
FT TRANSMEM 460 480 POTENTIAL.
FT DOMAIN 481 522 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 522 AA; 57085 MW; 075AB81E56CF33F7 CRC64;

```

```

Query Match 23.0%; Score 535; DB 1; Length 522;
Best Local Similarity 29.7%; Pred. No. 1e-28;
Matches 153; Conservative 90; Mismatches 167; Indels 106; Gaps 19;

```

```

OY 8 FAALGPISFGFALG-YSSP-AIPSLQRAAPPAALDDAA----- 46
DB 15 FTAVLGSGFQGDIGVINAPOEVIISHRYHLGV-LDDRATTIYDINGDTPLIYTPA 73
OY 47 -----SW-----FGAVVTL-----GAAAGVYL-----GGWLYDRAGRKSLLSVPE 84
DB 74 HTTPDAMEETEGSAHIVITMLMSLSVSSFAVGWAVSFFGGGLGDKLRIKAMLANSL 133
OY 85 VAGFAVITAAQ---DVMMLGGRLLTGLACGVASLAPVYISEIAPPAVAGLIGSCVOLM 141
DB 134 LTGALLMGCSKFGPAHALILAGRSVGLGCLISGLVPMYIGELIPTLRGALGTIHLA 193
OY 142 VVVGILLAYLAG--WVL-----EMRLAVLGCVPSRLMLLMCFMPEPRFL---LTQHR 192
DB 194 LVTGILISOIAGLSFTIGNODYWHILLGLSAPVALLQCLLFCESPRYLYLNEEVR 253
OY 193 QEATALRFLMGSBOGWD-----PIGAESFHL-LALLRQPGIYKPFILIGVSLMAFO 243
DB 254 AKK-SLKLRGTEDITRKINEMRKEKEASQKXSVIQLFTDPRVQRPIYVALMLHLAQ 312
OY 244 QLSGVNAVIFYAETIFEEAKFKDSSLASVYVGVYQVLFVAVALMDRAGRLLLVLSGV 303
DB 313 QPSGNGIFYYSTSIFQPAISQPYATIGVAINMIFTVASVLLVERAGRKTEF-LAGM 371
OY 304 VMVYSTSAFGAVFKITGCGPGNSHVALSAPVSAQVVDASVGL-----AMLVAGSMC- 355
DB 372 TGMFCAVFE-----MSLGLVLLDKFTWMSYSMTAI 402
OY 356 -LFIAGFVAGCPIPLMLMSEIFPLHVGAVATGICVLTNMLMAFLVTKFSSSLMEVLRPY 414
DB 403 FLVPSFPIGCPIDPMVAVAFEFSGRPRLALAAAFSNWNCNFIILALCFVIAFLGTY 462
OY 415 GAFWLASAFCIFSVLFTLFCVPETKGKTLQDTAHEF 450
DB 463 VFFLFAGVVLVFT-LTFPFKVPETKGSFDEIAAEF 497

```

RESULT 14

GTR2_MOUSE

ID CTR2-MOUSE STANDARD: PRT: 523 AA.
AC P14246;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE GLUCOSE TRANSPORTER TYPE 2, LIVER.
GN SLC2A2 OR GLUT2 OR GLUT-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C; TISSUE=Liver;
RC MEDLINE=90098776; PubMed=2602116;
RT Suzue K., Lodish H.F., Thorens B.;
RT "Sequence of the mouse liver glucose transporter.";
RT Nucleic Acids Res. 17:10099-10099(1989).
RN RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Liver;
RC MEDLINE=89366666; PubMed=2771649;
RA Asano T., Shibasaki Y., Lin J.L., Akanuma Y., Takaku F., Oka Y.;
RT "The nucleotide sequence of cDNA for a mouse liver-type glucose
RT transporter protein.";
RT Nucleic Acids Res. 17:6386-6386(1989).
RN RP SEQUENCE OF 384-496 FROM N.A.
RX MEDLINE=92111400; PubMed=1765007;
RA Hagan A., Heyner S., Charron M.J., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A., Thorens B., Schultz G.A.;
RT "Glucose transporter gene expression in early mouse embryos.";
RT Development 113:363-372(1991).
RN RP [4]
RP TISSUE SPECIFICITY.
RC STRAIN=C57BL/6;
RX MEDLINE=93170163; PubMed=1289053;
RA Smith D.E., Gridley T.;
RT "Differential screening of a PCR-generated mouse embryo cDNA library:
RT glucose transporters are differentially expressed in early
RT postimplantation mouse embryos.";
RT Development 116:555-561(1992).
CC -1- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY
CC MEMBRANE OF HEPATOCYTES & IS RESPONSIBLE FOR UPTAKE OF GLUCOSE BY
CC THE BETA CELLS. IT MAY COMPRISE PART OF THE GLUCOSE-SENSING
CC MECHANISM OF THE BETA CELL. IT MAY ALSO PARTICIPATE WITH THE
CC NC(+)/GLUCOSE COTRANSPORTER IN THE TRANSCELLULAR TRANSPORT OF
CC GLUCOSE IN THE SMALL INTESTINE & KIDNEY.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: IN EMBRYO, EXPRESSED IN ENDODERM LAYER OF
CC YOLK SAC AND LIVER PRIMORDIUM.
CC -1- SIMILARITY: HIGH TO OTHER GLUCOSE TRANSPORTERS.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X16986; CA34855.1; -
DR EMBL: X15684; CA33719.1; -
DR EMBL: S77926; AAB20847.1; -
DR PIR: S06920; S06920.
DR PIR: S05319; S05319.
DR MGI: 1095438; Slc2a2.
DR InterPro: IPR003663; Sugar_transport.
DR InterPro: IPR003662; sub_transport.
DR Pfam: PF00083; sugar_tr; 1.
DR PRINTS: PR00171; SUGRTRANSPORT.

DR PRINTS: PR00172; GLUCRTRANSPORT.
DR PRINTS: PR01191; GLUCRTRANSPORT2.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transporter; Transport; Glycoprotein;
KW Multigene family.
FT TRANSMEM 11 31 POTENTIAL.
FT DOMAIN 32 97 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 98 118 POTENTIAL.
FT TRANSMEM 127 147 POTENTIAL.
FT TRANSMEM 158 178 POTENTIAL.
FT TRANSMEM 187 207 POTENTIAL.
FT TRANSMEM 217 237 POTENTIAL.
FT DOMAIN 238 302 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 303 323 POTENTIAL.
FT TRANSMEM 338 358 POTENTIAL.
FT TRANSMEM 369 389 POTENTIAL.
FT TRANSMEM 400 420 POTENTIAL.
FT TRANSMEM 433 453 POTENTIAL.
FT TRANSMEM 461 481 POTENTIAL.
FT DOMAIN 482 523 CYTOPLASMIC (POTENTIAL).
FT CAROXYD 62 62 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 106 106 G -> D (IN REF. 2).
FT CONFLICT 181 181 N -> T (IN REF. 2).
FT CONFLICT 327 327 S -> T (IN REF. 2).
FT CONFLICT 351 351 L -> F (IN REF. 2).
FT CONFLICT 432 432 P -> S (IN REF. 1).
SQ SEQUENCE 523 AA; 57085 MW; 383F7808D73F0DFC CRC64;

Query Match 22.8%; Score 530.5; DB 1; Length 523;
Best Local Similarity 29.2%; Pred. No. 2e-28;
Matches 151; Conservative 90; Mismatches 168; Indels 109; Gaps 18;

QY 8 FAALGPTLSPFGALG-YSSP-----AIPSLGRA----- 35
DB 15 FTAIVLSFGYDGVINAPQEVIIISHRYHVLGVLDRKKAINDVAGTDTPLTVTPAY 74
QY 36 -PPARLDAAASWFGAVTL-----GAAAGVVL-----GWLVDRAGRKLSLLCSVP 83
DB 75 TTPAP-WDEETEGSAAHVTMLMSLVSSFAVGWASVFSGWLGDKIRIKAMLAANSL 133
QY 84 FVAGFAVITTAO---DWMMLGRLTGLAGVASLVAPVYISLAFYAVAGLSCVQL 140
DB 134 SLTGALLMGCKFGPAHALIITAGRSVSLYGLLSGLVPMYIGELAPNTLGAULTLHQL 193
QY 141 MVVVGILLAYLAG--WVL-----EWRMLAVLGCVPSSLMLLMCFMPTPRFLTLQHRROE 194
DB 194 ALVIGILTSQIALGSLIFLIGNODHMHILGLSAVALLQCLLLCPESPRLYI--KLE 251
QY 195 AI-----ALRFLMGSEQGMED-----PTGADQSFALLLRQPGIYK-PFTIGVSLMA 241
DB 252 EVRAKSKLRKRGEDVTKDINEMKKKEEASTOKKSVIOLFLFDANYROPILVALMLHM 311
QY 242 FOOLSGVNAVVFVAETTFEAKFKDSSIASVAVGVIOFLFAVALALMDRGRLLVLVS 301
DB 312 AQQSGINGIFYYTSTIFQAGISQIPYATIGVCAIMILLTASVLLVERAGRRTL-LT 370
QY 302 GVWVVFSTSAFGAVFKLTGGPGNSSHVAISAPVASQVPDASVGL-----AWLAVGSMC 355
DB 371 GMIQMFCTIF-----MSVGLVLLDRFAMMSYVSMT 401
QY 356 ---LFIQPAVNGPIRPMWLMSEIFPLHVGAVATGICVNLNMAFLVTRFSSLMELVR 412
DB 402 AILFLVSEFFELGPGPLFWVAEVEFSSGPPRTALALAAFNWCVNCFVIALCFYIADFLG 461
QY 413 PYGAFWLASAFCLFSVLTFLFCVETKGTLEQITAE 450
DB 462 PYVEFLFAGVAVIVPT-LFTTFKVPETKGSFEELAEF 498
RESULT 15
HUP3_CHLKE

ID HUP3_CHLKE STANDARD: PRT: 534 AA.
AC Q39525:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE H(+)/HEXOSE COTRANSPORTER 3.
GN HUP3.
OS Chlorella kessleri.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales.
OX NCBI_TaxID=3074;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95175627; PubMed=7870840;
RA Stadler R., Wolf K., Hilgarch C., Tanner W., Sauer N.K.;
RT "Subcellular localization of the inducible Chlorella HUP1 monosaccharide-H⁺ symporter and cloning of a Co-induced galactose-H⁺ symporter";
RL Plant Physiol. 107:33-41(1995).
CC -1- FUNCTION: ACTIVE UPTAKE OF HEXOSES.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X75440; CAA53192.1;
DR InterPro: IPR003663; Sugar_transport.
DR InterPro: IPR003662; sub_transport.
DR Pfam: PF00083; sugar_tf_1.
DR PRINTS: PR00171; SUGARTRANSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Transport; Sugar transport; Symport; Repeat;
KW Multigene family.
FT TRANSMEM 21 41 POTENTIAL.
FT TRANSMEM 87 107 POTENTIAL.
FT TRANSMEM 114 134 POTENTIAL.
FT TRANSMEM 137 157 POTENTIAL.
FT TRANSMEM 172 192 POTENTIAL.
FT TRANSMEM 205 225 POTENTIAL.
FT TRANSMEM 306 326 POTENTIAL.
FT TRANSMEM 329 349 POTENTIAL.
FT TRANSMEM 353 373 POTENTIAL.
FT TRANSMEM 395 415 POTENTIAL.
FT TRANSMEM 431 451 POTENTIAL.
FT TRANSMEM 456 476 POTENTIAL.
SO SEQUENCE 534 AA; 57772 MW; 52E03715F3726017 CRC64;

Query Match 22.8%; Score 530.5; DB 1: Length 534;
Best Local Similarity 30.9%; Pred. No. 2,1e-28;
Matches 151; Conservative 77; Mismatches 195; Indels 65; Gaps 14;

OY 3 VFLAFAAALGLSFCFALGYS---SPAIPSLOR-----AAPAPRLD 42
DB 26 VLVAVLVAAAG---GMLGVDNGVGTGVASMEQERKFPDYERKKQDIVETSPCYTD 81
OY 43 DAASMPGAVVTLGAAGVGVLGMLVDRAGRKLSLLCSVPFVAGFAVITA--AQDVWMLL 101
DB 82 NPKLDLPFVSLFAGLISCFISAMITRNWGRKASKMGICGIFFIAGGLVNAFQDITAMLI 141
OY 102 GRLTLTGACGVASLVAPVYISEIAYPAVRGLLGSCVQLMNVVGTLLAVIA-----GWVL 156
DB 142 VGRVLLGFGVGISGVVPOYLSEVAPFSHRGMLNIGYQLFTVITGILIAQLVNVGVNMDN 201
OY 157 EWRWLAIVAGVPPSLMLLMCFMPETPRLL-----TQHRQREAIALRPLMGSEQGMED-- 210
DB 202 GWRSLGLAAVGLITLLGAIVLPESPFLVERKRTDGRILILEKRTSHVEAFADIV 261

OY 211 -----PPIGAESFHLALRQ--PGIYKPIIGVLSMAFOQLSGVNAVMEFAETFEEA 262
DB 262 AAVEIARPIITMRQSWRSLPFRRYMPQLTTSFVI-----QFQOQFGINALIIVVPLFSSL 317
OY 263 KFKDSS--LASVVGVIQVLTAVAAALIMDRAGRLLLVISGVVWVSTSAFGAYFKLTQ 320
DB 318 GSASSAALLMTVVVGAENVSTMTAVLLSDKFGRRPFLIEGRTCCLAMIAAGITLGVEF 377
OY 321 GPGNSSHVAIISAPVSAQPYDASVGLAMIAVGSNCLFIAGFVAGMGP1PMLTMSFIFPLH 380
DB 378 GQGTED-----LPHVSAQV---LAV--ICITFAGFAMSGMPGWLIPSEIFPLE 423
OY 381 VKGATGICVLTNMLAFLVTKFESSLMELVRPYGAFWLASAFICFSVLTFLPCVPETKG 440
DB 424 TRAGTAVAVMGNFLPFSFVIGQAFVSMICAMK--FGVLFIFAGMLVIVNLCALFILPETKG 482
OY 441 KYLEQITA 448
DB 483 VPIERVQA 490

Search completed: February 13, 2002, 22:06:05
Job time: 861 sec

Thu Feb 14 07:44:56 2002

us-09-516-493-7.rsp

Page 14

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 21:50:19 ; Search time 136.8 Seconds

(without alignments)
484,367 Million cell updates/sec

Title: US-09-516-493-7

Perfect score: 2324
Sequence: 1 RRVFLAFAAALGPLSGFA.....CVPEKCKTLEQITAIHFGR 453

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2301.5	99.0	477	4 Q9NY64	Q9ny64 homo sapien
2	2298.5	98.9	477	4 Q9NSC4	Q9ns44 homo sapien
3	2057	88.5	477	11 Q9JJZ1	Q9jjz1 rattus norv
4	2050.5	88.2	477	11 Q9JJZ4	Q9jjz4 mus musculu
5	2047.5	88.1	477	11 Q9JJZ3	Q9jjz3 mus musculu
6	2041.5	87.8	477	11 Q9JJZ0	Q9jjz0 mus musculu
7	2040.5	87.8	479	11 Q9JMA6	Q9jma6 rattus norv
8	927	33.9	507	4 Q9UGQ3	Q9ugq3 homo sapien
9	780.5	33.6	539	4 Q9VU17	Q9vu17 drosophila
10	777	33.4	489	5 Q9V609	Q9v609 drosophila
11	753	32.4	433	5 Q9V610	Q9v610 drosophila
12	730	31.4	515	10 Q9LMA8	Q9lma8 arabidopsis
13	725.5	31.2	487	10 Q9FRJ3	Q9frj3 arabidopsis
14	717.5	30.9	465	5 Q9V848	Q9v848 beta vulgar
15	717.5	30.9	490	10 Q9J416	Q9j416 oryza sativ
16	708.5	30.5	501	10 Q9JHW3	Q9jhw3 oryza sativ
17	692	29.8	497	5 Q9VNU1	Q9vnu1 drosophila
18	656	28.2	444	5 Q9VNG0	Q9vng0 drosophila
19	637.5	27.4	491	5 Q9VNU2	Q9vnu2 drosophila

20	636.5	27.4	463	10 Q9MAA4	Q9maa4 arabidopsis
21	577.5	24.8	521	10 Q22848	Q22848 arabidopsis
22	572	24.6	496	10 Q65799	Q65799 arabidopsis
23	555.5	23.9	580	10 Q9C757	Q9c757 arabidopsis
24	554	23.8	510	10 Q65322	Q65322 petunia hyb
25	550	23.7	457	2 P96742	P96742 bacillus su
26	549	23.6	534	10 Q9LLE1	Q9lle1 nicotiana t
27	548	23.6	515	10 Q9LUD8	Q9lud8 arabidopsis
28	545	23.5	464	2 P96710	P96710 bacillus su
29	544.5	23.4	477	10 Q9SCW9	Q9scw9 arabidopsis
30	544.5	23.4	558	10 Q9RFP2	Q9rfp2 arabidopsis
31	544	23.4	470	10 Q9LLE0	Q9lle0 solanum tub
32	543	23.4	542	10 Q9LUD9	Q9lud9 zea mays (m
33	542	23.3	560	10 Q9LFI3	Q9lfi3 arabidopsis
34	538.5	23.2	473	2 Q05182	Q05182 bacillus me
35	538	23.1	551	10 Q9LLE2	Q9lle2 spinacia ol
36	537.5	23.1	523	11 Q9DBA7	Q9dba7 mus musculu
37	537	23.1	502	10 Q9FRP7	Q9frp7 oryza sativ
38	537	23.1	555	10 Q9FRY8	Q9fry8 mesembryant
39	536	23.1	461	10 Q9M100	Q9m100 arabidopsis
40	534.5	23.0	518	10 Q40373	Q40373 medicago tr
41	534.5	23.0	523	10 Q06312	Q06312 nicotiana t
42	534	23.0	526	10 Q81501	Q81501 arabidopsis
43	530.5	22.8	518	10 Q9FRF5	Q9frf5 oryza sativ
44	530	22.8	581	10 Q9LKH2	Q9lkh2 mesembryant
45	525	22.6	507	10 Q81492	Q81492 arabidopsis

ALIGNMENTS

```
RESULT 1
Q9NY64 PRELIMINARY; PRT; 477 AA.
ID Q9NY64:
AC Q9NY64:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE GLUCOSE TRANSPORTER.
GN GLUTXL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20138191; PubMed-10671487;
RA Ibbertson M.R., Uldry M.A., Thorens B.;
RT "GLUTXL: A novel mammalian glucose transporter expressed in the
RT central nervous system and insulin-sensitive tissues.";
RL J. Biol. Chem. 275:4607-4612(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL: AJ245937; CAB75702.1; -.
DR InterPro: IPR003662; sub.transporter.
DR InterPro: IPR003663; Sugar transporter.
DR Pfam: PF00083; sugar_tr.1.
DR PRINTS: PR00171; SUGRTNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane.
SQ SEQUENCE 477 AA; 50858 MW; 39DA9E2DBFF8D3E2 CRC64;
```

Query Match 99.0%; Score 2301.5; DB 4; Length 477;

Best Local Similarity 99.1%; Pred. No. 1,1e-122;
Matches 450; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

```
QY 1 RRVFLAFAAALGPLSGFALGYSSPAIPSLQRAAPPAPRLDDAAASFAGAVTTGAAG 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 24 RRVFLAFAAALGPLSTGFALGYSSPAIPSLQRAAPPAPRLDDAAASFAGAVTTGAAG 83
QY 61 GVLGWLVDRAGRLSLLLCSVPVAGFAVITAQDVWMLLGGRLLTGLACGVASLVAPV 120
```

```

|||||
Db 84 GVLGWLVDRAGRKLSLLCGVPEVAGFAVTTAODVWMLLGGRLTLGLAGVASLVA PV 143
|||
QY 121 YISEIATPAVARGLLGSCVQMLVNVVGIILLAYLAGVLEMRMLAVLGCVPSPMLLLMCFMP 180
|||
Db 144 YISEIATPAVARGLLGSCVQMLVNVVGIILLAYLAGVLEMRMLAVLGCVPSPMLLLMCFMP 203
|||
QY 181 ETPEFLLTQHRROBAT-ALRFLMGSEOGMEDPPIGAEOSEPHALLRPGIKYKFFIIGVSL 239
|||
Db 204 ETPEFLLTQHRROBAMALRFLMGSEOGMEDPPIGAEOSEPHALLRPGIKYKFFIIGVSL 263
|||
QY 240 MAFQOLSGVNAVMFYAETTFEEAKFKDSSLASVVGVIOVLTFAVALIMDRAGRLLLV 299
|||
Db 264 MAFQOLSGVNAVMFYAETTFEEAKFKDSSLASVVGVIOVLTFAVALIMDRAGRLLLV 323
|||
QY 300 LSGVWVFSTSAFGAYFKLQGGPGNSSHVAISAPVSAQPVDAVGLAMVLAVGSMCLFIA 359
|||
Db 324 LSGVWVFSTSAFGAYFKLQGGPGNSSHVAISAPVSAQPVDAVGLAMVLAVGSMCLFIA 383
|||
QY 360 GFAVGMGPFWLMLMSEIFPLHVKGVATGICVLTNWMLMAFLVTKFESSLMELVRPYGAFWL 419
|||
Db 384 GFAVGMGPFWLMLMSEIFPLHVKGVATGICVLTNWMLMAFLVTKFESSLMELVRPYGAFWL 443
|||
QY 420 ASAFICFSVLTFLCVPETKGTLEQITTAHEGR 453
|||
Db 444 ASAFICFSVLTFLCVPETKGTLEQITTAHEGR 477

```

```

RESULT 2
Q9NSC4 PRELIMINARY; PRT; 477 AA.
ID Q9NSC4;
AC Q9NSC4;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE GLUCOSE TRANSPORTER 8.
GN GLUT8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=20283667; Pubmed=10821868;
RA Doege H., Schuermann A., Bahrenberg C., Brauers A., Joost H.G.;
RT "Glut8: A novel member of the sugar transport facilitator family with
RT glucose transport activity."
RL J. Biol. Chem. 275:16275-16280(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL: Y17801; CAB89809.1; -.
DR InterPro: IPR003662; sub_transporter.
DR InterPro: IPR003663; Sugar_transporter.
DR Pfam: PF00083; sugar_tr.1.
DR PRINTS: PR00171; SUGRTRNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane.
SQ SEQUENCE 477 AA; 50792 MW; 0B480F94B40AE76 CRC64;

```

```

Query Match 98.9%; Score 2298.5; DB 4; Length 477;
Best Local Similarity 99.1%; Pred. No. 1,7e-122;
Matches 450; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
QY 1 RRVLAAFAAALGPLSGFALGYSSPAIPSLQRAAPAPRLDDAAASWFGAVVTLGAAG 60
|||
Db 24 RRVLAAFAAALGPLSGFALGYSSPAIPSLQRAAPAPRLDDAAASWFGAVVTLGAAG 83
|||
QY 61 GVLGWLVDRAGRKLSLLCGVPEVAGFAVTTAODVWMLLGGRLTLGLAGVASLVA PV 120
|||
Db 84 GVLGWLVDRAGRKLSLLCGVPEVAGFAVTTAODVWMLLGGRLTLGLAGVASLVA PV 143

```

```

QY 121 YISEIATPAVARGLLGSCVQMLVNVVGIILLAYLAGVLEMRMLAVLGCVPSPMLLLMCFMP 180
|||
Db 144 YISEIATPAVARGLLGSCVQMLVNVVGIILLAYLAGVLEMRMLAVLGCVPSPMLLLMCFMP 203
|||
QY 181 ETPEFLLTQHRROBAT-ALRFLMGSEOGMEDPPIGAEOSEPHALLRPGIKYKFFIIGVSL 239
|||
Db 204 ETPEFLLTQHRROBAMALRFLMGSEOGMEDPPIGAEOSEPHALLRPGIKYKFFIIGVSL 263
|||
QY 240 MAFQOLSGVNAVMFYAETTFEEAKFKDSSLASVVGVIOVLTFAVALIMDRAGRLLLV 299
|||
Db 264 MAFQOLSGVNAVMFYAETTFEEAKFKDSSLASVVGVIOVLTFAVALIMDRAGRLLLV 323
|||
QY 300 LSGVWVFSTSAFGAYFKLQGGPGNSSHVAISAPVSAQPVDAVGLAMVLAVGSMCLFIA 359
|||
Db 324 LSGVWVFSTSAFGAYFKLQGGPGNSSHVAISAPVSAQPVDAVGLAMVLAVGSMCLFIA 383
|||
QY 360 GFAVGMGPFWLMLMSEIFPLHVKGVATGICVLTNWMLMAFLVTKFESSLMELVRPYGAFWL 419
|||
Db 384 GFAVGMGPFWLMLMSEIFPLHVKGVATGICVLTNWMLMAFLVTKFESSLMELVRPYGAFWL 443
|||
QY 420 ASAFICFSVLTFLCVPETKGTLEQITTAHEGR 453
|||
Db 444 ASAFICFSVLTFLCVPETKGTLEQITTAHEGR 477

```

```

RESULT 3
Q9JUZ1 PRELIMINARY; PRT; 478 AA.
ID Q9JUZ1;
AC Q9JUZ1;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE GLUCOSE TRANSPORTER.
GN GLUTX1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2018191; Pubmed=10671487;
RA Ibbertson M.R., Uldry M.A., Thorens B.;
RT "Glutx1: A novel mammalian glucose transporter expressed in the
RT central nervous system and insulin-sensitive tissues."
RL J. Biol. Chem. 275:4607-4612(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL: AJ245935; CAB75729.1; -.
DR InterPro: IPR003662; sub_transporter.
DR InterPro: IPR003663; Sugar_transporter.
DR Pfam: PF00083; sugar_tr.1.
DR PRINTS: PR00171; SUGRTRNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane.
SQ SEQUENCE 478 AA; 51458 MW; 95841FC1F18C9BE9 CRC64;

```

```

Query Match 88.5%; Score 2057; DB 11; Length 478;
Best Local Similarity 85.7%; Pred. No. 7.1e-109;
Matches 390; Conservative 32; Mismatches 31; Indels 2; Gaps 2;
QY 1 RRVLAAFAAALGPLSGFALGYSSPAIPSLQRAAPAPRLDDAAASWFGAVVTLGAAG 60
|||
Db 24 RRVLAAFAAALGPLSGFALGYSSPAIPSLQRAAPAPRLDDAAASWFGAVVTLGAAG 83
|||
QY 61 GVLGWLVDRAGRKLSLLCGVPEVAGFAVTTAODVWMLLGGRLTLGLAGVASLVA PV 120
|||
Db 84 GVLGWLVDRAGRKLSLLCGVPEVAGFAVTTAODVWMLLGGRLTLGLAGVASLVA PV 143
|||
QY 121 YISEIATPAVARGLLGSCVQMLVNVVGIILLAYLAGVLEMRMLAVLGCVPSPMLLLMCFMP 180
|||

```


Db	144	YISLAPAVAGLLGSCVOLNVYIGILLAIVAGAVLEMRHLAVLGCPPLMLLMQYMP	203
Qy	181	ETPRELLTOHROEAT-ALRFMCSBOGMDPRIGAE-QSEHLALLKQCPYKPPILIGVS	238
Db	204	ETPRELLTOHOYOAMALRLFMCSBOGMDPRPGAEHQGFOLAMLRPGVHKPLILIGIC	263
Qy	239	LMARQOLSGVAWVEFYAETTFEEKAFKDDSLASVYVGVIOVLTAVALLIMDRAGRRL	298
Db	264	LMVQOLSGVAWIMFYANTTFEEKAFKDDSLASVTVGIQVLTAVALLIMDRAGRRL	323
Qy	299	VLSGVWVEFSTSAFGATFKLTGGPGNSSHVAISAPVSAQPVDAVGLAMLAVGSMCLT	358
Db	324	ALSGVIMVFSMAFGATFYFKLTQSPSSNHHGLVPRISAPRADVHGLAMLAVGSMCLT	383
Qy	359	AGFAVGMQPIEMLLMSELFLPHVKGAVTGVCVLTNMLMAFLVTKRFSLEVLIRPGAFM	418
Db	384	AGFAVGMQPIEMLLMSELFLPHIKGAVTGVCVLTNNWMAFLVTKRFSIMELLIRPGAFM	443
Qy	419	LASAFCTFSVLTFLCYVPETKGLTLEDTITAHFEGR	453
Db	444	LTAAFCTLSVLTFLTVEVPETKGRGLTITAHFEGR	478

RESULT	4			
09JUP4				
ID	09JUP4	PRELIMINARY:	PRT:	477 AA.
AC	09JUP4;			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, last annotation update)			
DE	GLUCOSE TRANSPORTER 8.			
GN	GLUTX1 OR GLUT8.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheraia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_Taxid=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=TESTIS;			
RX	MEDLINE=20281667; PubMed=10821868;			
RA	Doerge H., Schuermann A., Bahnenberg C., Brauers A., Joost H.G.;			
RT	"GLUT8: A novel member of the sugar transport facilitator family with			
RT	glucose transport activity."			
RL	J. Biol. Chem. 275:16275-16280(2000).			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).			
CC	-1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.			
DR	EMBL, Y17802; CAB89815.1; -.			
DR	MCD; MGI:1860103; Glutx1.			
DR	InterPro; IPR003662; sub1.transporter.			
DR	InterPro; IPR003663; Sugar.transporter.			
DR	Pfam; PF000083; sugar tr; 1.			
DR	PRINTS; PR00171; SUGTRNSPORT.			
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN-2.			
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.			
DR	Transmembrane.			
Q0	SEQUENCE 477 AA: 51507 MW: 59A985B52FADF478 CRC64;			

[illegible]

QY	181	ETPRLLTQHRROEPT -ALRFLMSQEQMDPPRIGAE -QSFHLALLKQPGIYKPFIIIGS	238
Db	204	ETPRRLTLQHQYQEMALRLKMSSEBMEHPYPGAEGHQGLALLRRPGIYKRLIIIGIS	265
QY	239	LMAFQOLSGVNAVMEYETFEFEAKFKDSSLASVVGVIQVLPMAVALIMDRAGRLL	299
Db	264	LMVFOOLSGVNAVIMFYANSIFEFEAKFKDSSLASVYGVLIQVLPFAVALIMDRAGRLL	323
QY	299	VLSTGVVMEFNSARCAVYKFLQGGPGNSHVAISAPYSAQVVDASVGLAMLVGSMCLFI	358
Db	324	ALISGVIMFESMAEFTYKFLQSLPSNSHGL -VPLAAEVDVDQVGLAMLVGSMCLFI	382
QY	359	AGFAVGMCPIMLLMSELPLTHVKGAVATGIVLTNNMLAFLVTKFSSLMEDLVLPYGAFW	418
Db	383	AGFAVGMCPIMLLMSELPLTHVKGAVATGIVLTNNMFAFLVTKFSSVMEDLPRPYGAFW	442
QY	419	LASAFICISVLEFTLYFCVDETNGKLTLEQYTNAAEEGR	453
Db	443	LTAACFALSVLEFTLYVPEPTNGRLTLEQYTNAAEEGR	477

SEQUENCE	477 AA	51523 MW	A3753FB34E452F9A	CRC64
DR	PROSITE, PS00217; SUGAR_TRANSPORT_2; 1.			
DR	PROSITE, PS00171; SUGAR_TRANSPORT_1; UNKNOWN_1.			
DR	PRINTS, PR00121; SUGAR_TRANSPORT.			
DR	PIfam; PF00083; sugar_tf; 1.			
DR	InterPro; IPR003663; Sugar_transporter.			
DR	InterPro; IPR003662; sub_transporter.			
DR	MGP; MGT1860103; GLUTX1.			
DR	EMBL; AF232061; AAF78366.1; -			
CC	-1- SIMILARITY: TO SUGAR_TRANSPORTER FAMILY.			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).			
RL	Proc. Natl. Acad. Sci. U.S.A. 97:7313-7318(2000).			
RT	"GLUT8 is a glucose transporter responsible for insulin-stimulated glucose uptake in the blastocyst."			
RA	McKnight R.A., Mueckler M., Devaskar S.U., Moley K.H.;			
RA	Carayannopoulos M.O., Chi M.M.Y., Cui Y., Pingsterhaus J.M.,			
RC	MEHLIN=20313023; PubMed=10860996;			
RP	SEQUENCE FROM N.A.			
RP	STRAIN=129;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
GN	GLUTX1.			
DE	GLUCOSE TRANSPORTER GLUT8.			
DT	01-OCT-2000 (TRENBLREL, 15, Last sequence update)			
DT	01-OCT-2000 (TRENBLREL, 15, Last sequence update)			
DT	01-JUN-2001 (TRENBLREL, 17, Last annotation update)			
ID	09JIF3; PRELIMINARY; PRT; 477 AA.			
AC	09JIF3;			
AC	01-OCT-2000 (TRENBLREL, 15, Created)			
AC	01-OCT-2000 (TRENBLREL, 15, Last sequence update)			
AC	01-JUN-2001 (TRENBLREL, 17, Last annotation update)			
AC	GLUCOSE TRANSPORTER GLUT8.			
AC	GLUTX1.			
AC	Mus musculus (Mouse).			
AC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AC	NCBI_TaxID=10090;			
AC	[1]			
AC	SEQUENCE FROM N.A.			
AC	STRAIN=129;			
AC	MEHLIN=20313023; PubMed=10860996;			
AC	Carayannopoulos M.O., Chi M.M.Y., Cui Y., Pingsterhaus J.M.,			
AC	McKnight R.A., Mueckler M., Devaskar S.U., Moley K.H.;			
AC	"GLUT8 is a glucose transporter responsible for insulin-stimulated glucose uptake in the blastocyst."			
AC	Proc. Natl. Acad. Sci. U.S.A. 97:7313-7318(2000).			
AC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).			
AC	-1- SIMILARITY: TO SUGAR_TRANSPORTER FAMILY.			
AC	EMBL; AF232061; AAF78366.1; -			
AC	MGP; MGT1860103; GLUTX1.			
AC	InterPro; IPR003662; sub_transporter.			
AC	InterPro; IPR003663; Sugar_transporter.			
AC	PIfam; PF00083; sugar_tf; 1.			
AC	PRINTS, PR00121; SUGAR_TRANSPORT.			
AC	PROSITE, PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.			
AC	PROSITE, PS00217; SUGAR_TRANSPORT_2; 1.			
AC	Transmembrane.			
AC	SEQUENCE 477 AA; 51523 MW; A3753FB34E452F9A CRC64;			

	Query Match	88.1%	Score 2047.5	DB 11	length 477
	Best Local Similarity	86.4%	Pred. No. 2,4e-108		
	Matches 193	Conservative 30	Mismatches 29	Indels 3	Gaps 3
QY	1 RRVPLAFAAALGPIISFGALGYSSPAIPISLORAPAPRLDADAASWGCAYVTLGCAAG	60			
DB	24 RRVPLASPAALGPIISFGALGYSSPAIPISLRTAPALRLGDMAASWGCAYVTLGCAAG	83			
QY	61 GVLGMLVDRAAGKLSLLICSVPFVAGFAVITAADVWMLLGGRLLTGLACGVASIVAPY	120			
DB	84 GILGMLLDRCGRKLSLLICTVPEFTGFAVITAARDVWMLLGGRLLTGLACGVASIVAPY	143			
QY	121 YISEIATPAVRKLLGSCVOYLMMVVGILLATLACGVLEHMLAVLCGPPSLMLLMCFMP	180			
DB	144 YISEIATPAVRKLLGSCVOYLMMVVTGLIAYVGVLEHMLAVLCGPPSLMLLMCFMP	203			

Oy	181	ETPFFLLTOHROKAI-ALRFIMSECGMEPPIGAE-QSEFHLALLRQPIYKPTIIIGVS	238
Db	204	ETPRLILTHOYQEMALMLKMSSECKMEPPGAGHOGFOLALLRPPQITVPLIIIGIS	263
Oy	239	LMAFQOLSGVNAVMAFVAETIEEEKFKDDSSLASVVGVIQVLTFAVAALLMDRAGRRLTL	298
Db	264	LMVEFQOLSGVNAIMFVANSIIEEPAKFDDSSIASVTVGIIQVLTFAVALLMDRAGRRLTL	323
Oy	299	VLSGVMAVFFTSARGAVFEXKLQGGPRGSSHVAISAPVSAQPVDAVGLATLAVGSMCLFI	358
Db	324	ALSGVIMFMSAGFYTKLQSLPSSMSHGL-VPIAAEVDVDQVGLAALVGSMLFI	382
Oy	359	AGFAVGMCPIDWMLMSEIFPLHNVGAVATGICVLTNWIMAEFLYTKKESSLMEVLRYPGAFW	418
Db	383	AGFAVGMCPIDWMLMSEIFPLHNVGAVATGICVLTNWEMAEFLYTKKESSVWEMLRYPGAFW	442
Oy	419	LASAFCTFSVLPFTLCFVPEYTKGKLTLEDTTHAFEBR	453
Db	443	LTAAFCAISLVLPFTLVVPEYTKGRRLDEVYTHAFEBR	477

```

RESULT      6
09JUZ0
ID      09JUZ0      PRELIMINARY:      PRT:      477 AA.
AC      09JUZ0;
DT      01-OCT-2000 (Tremblrel, 15, Created)
DT      01-OCT-2000 (Tremblrel, 15, Last sequence update)
DT      01-JUN-2001 (Tremblrel, 17, Last annotation update)
DE      GLUCOSE TRANSPORTER.
GN      GLUTX1.
OS      Mus musculus (mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20138191; PubMed=10671487;
RT      Ibersen M.R., Uldry M.A., Thorens B.;
RT      "GLUTX1: A novel mammalian glucose transporter expressed in the
RT      central nervous system and insulin-sensitive tissues.";
RL      J. Biol. Chem. 275:4607-4612(2000).
CC      1. SUBCELLULAR LOCATION: INTERAL MEMBRANE PROTEIN (BY SIMILARITY)
CC      -1. SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
CC      EMBL: AJ245936; CAB75719.1; -.
CC      MGI: 1860103; Glutx1.
DR      InterPro: IPR003662; sub.transporter.
DR      InterPro: IPR003663; Sugar transporter.
DR      Pfam: PF00063; sugar_tr_1.
DR      PRINTS: PR00171; SUGCTRNSPORT.
DR      PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
DR      PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
DR      Transmembrane.
FW      SEQUENCE      477 AA;      51578 MW;      14FF23DB9060DAF1 CRC64;

```

	Query Match	87.88%	Score 2041.5	DB 11	Length 477
	Best Local Similarity	85.98%	Pred. No. 5.3e-108		
	Matches	391	Conservative	32	Mismatches 29; Indels 3; Gaps
Oy	1 RRVFLAAFAAALGPLSEFCFALGYSPPALPSLQORAPAPARLDDDAASMGAVVTLGAAG	60			
Dd	24 RRVLASTAAALGLPLNPFALGYSPPALPSLRRTAPALRLGDMAASWFGAVVTLLGAAG	83			
-Oy	61 GVLGMLVDNRGRKLSLLLCSPYVAGVATTAAODVMMLGGRLLTGLACVASLAAPV	120			
Dd	84 GILGWMLEDRSGRKLSLLLCCTVPFTGVGAVTTAADVMMLGGRLLTGLACVASLAAPV	143			
Oy	121 YISLEYAVRGELGSCVOLMWWVGITLAYLAGVWLEMRRLAYVCVPSTLMMLMCMP	180			
Dd	144 YISELAYPAVRKLGLASCVOLMWVGITLAYIAGVWLEMRRLAYVCVPTLMMLMKYMP	203			
Oy	181 ETPRFLTQHRRKEAI-LAFRLMSGEQGMEDPDTCAE-QSFHALLRPQGIYKPFIIGVS	238			

Db	204	ETPRRLTLHQVQGEAMALRLTLMGSEEGWEPVGAHEHGFQDLALRRGIYKPLIGIS	263
Qy	239	LAAFDQLSGVNAVMFYACTIFEEAKKFDSSLASVYGVQVLYFTVAALLMDRAGRLL	298
Db	264	LAMFDQLSGVNAIMFYANISIFEAKKFDSSLASVYGVQVLYFTVAALLMDRAGRLL	323
Qy	299	VLSGVMMFTSTAFGAYKRLKLOGGPGMSSHVAISAPVASQPDASVGLAMLAVGSKCLFI	358
Db	324	ALSGVIMFTSMKAFGYTKRLQSLSPSSSHVGL-VPIAEPVDVGVGLAMLAVGSKCLFI	382
Qy	359	AGFAVMGCIPIPLMLSEIFPLHYKGVATGICVLTNNLMALFVTKREFSSLMELRYGAFW	418
Db	383	AGFAVMGCIPIPLMLSEIFPLHYKGVATGICVLTNNMFMALVTKREFNSVMELRYGAFW	442
Qy	419	LASAFCTISVLTFLCVPEVTGKLTBOIT7AHFECR	453
Db	443	L7AFAFCAISVLTFLVPEVTGKLTBOIT7AHFECR	477

RESULT	7		
Q9JMA6		PRELIMINARY:	PRT; 479 AA.
ID	Q9JMA6		
AC	Q9JMA6;		
DT	01-OCT-2000 (TREMblrel. 15, Created)		
DT	01-OCT-2000 (TREMblrel. 15, Last sequence update)		
DT	01-JUN-2001 (TREMblrel. 17, Last annotation update)		
DE	GLUCOSE TRANSPORTER 8.		
GN	GLUT8.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RT	Ishidaishi K.;		
RA	Molecular cloning of a new putative glucose transporter.";		
RC	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.		
RL	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).		
CC	-1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.		
DR	EMBL: AB033418; BAA94383.1; -		
DR	InterPro: IPR003662; sub_transporter.		
DR	InterPro: IPR003663; Sugar_transporter.		
DR	Pfam: PF00083; sugar_tr.1.		
DR	PRINTS: PR00171; SUGRTNSPORT.		
DR	PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.		
DR	PROSITE: PS00217; SUGAR_TRANSPORT_2.1.		
DR	Transmembrane.		
QJ	SEQUENCE 479 AA: 51601 MW; EAB2B67B8CDD2PFC CRC64;		

[illegible]

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL: AE003540; AAF49874.1; -
 DR FlyBase: FBgn0036316; CG10960.
 DR InterPro: IPR003662; sub_transporter.
 DR InterPro: IPR003663; Sugar_transporter.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGARTRANSPORT.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane.
 SQ SEQUENCE 539 AA; 58413 MW; A2F481ADB676BBD CRC64;

Query Match 33.6%; Score 780.5; DB 5; Length 539;
 Best Local Similarity 37.6%; Pred. No. 7,9e-37;
 Matches 174; Conservative 90; Mismatches 166; Indels 33; Gaps 6;

4 FLAFAAALGPLSGFALGYSSPAIPSL-ORAAPAPRLDDAASWFGAVVTGGAAGGV 62
 84 YVAALAAAGAFAGATVLTGMTSPAETIVDRGEGYDPPDKDPSWGSANTLGAACVCI 143
 63 LGGWLVDRAGRKLSLLCSVPFAVAVTAADVWMLLGRLLTGLACGVALAPVYI 122
 144 PIGFLINMGRKMTMLFLVLPFLIGMTMLIWAIVNSMLYASRFLTGAGCAFVATADMT 203
 123 SELAYPAVRGLGSCVOLAMVVGILLAYLAGWLEEMWMLAVLGVPSMLMLMCFEPET 182
 204 GEIAQKEIRKGLSEFGLMTLITIGLFLYAVAGAGKIFMELITICILPLIGALFFEMES 263
 183 PRELLTOHRROEAL-ARFLMGSEOGWE-----DPIGABOSFHALLRPGIYKP 232
 264 PTYLVASDRBENNAIKSIQWLRGKDYDEPELAELRETRDKANKVWMAALNRPYRKA 323
 233 FIIGVSLMAFOQLSGVNAWVFATFEAKFK-DSSLASVVGVIQVLEFTAALIMDR 291
 324 LAISMGIMFPOQGVGNAVFAYASRIEFLNANTGIEAEMATILIGIMOVAVFESTLVVDK 383
 292 AGRRLILVLGSGVNVFSTSAFGAFLKLGQPGSSHVATISAPVSADPVAS--VGLANL 349
 384 LGRRLILLAGSISMAISTTAIGVFPL-----QKDAQVVSIGWL 424
 350 AVGSMCLFIAGFAGWGPIMLMSEIFPLHVKGATGICVLNMLAFLVTEFSSIME 409
 425 PVASLCLEFIIMFSTIGCPVPMWLMGELFATDIDKGFAGSLAGTSMWLLAFVTKTFVALND 484
 410 VLRYGAFWLASAFICFVSFLTELCVPEPKGKLTLEQITAHFEG 452
 485 GLIGIGTFEMLAGLTIVGVIFVYFAVPEPKGKSLNEIQDELAG 527

RESULT 10
 09V609 PRELIMINARY; PRT; 489 AA.
 AC 09V609;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 *DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE CG7801 PROTEIN.
 GN CG7801.
 CS *Drosophila melanogaster* (Fruit fly).
 *OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephyroidea; *Drosophilidae*; *Drosophila*.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt E.G., Nelson C.R., Miklos G.L.G.,
 RA Adair J.E., Agdayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mierke D., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclt J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL: AE003825; AAF58631.1; -
 DR FlyBase: FBgn0033643; CG7801.
 DR InterPro: IPR003662; sub_transporter.
 DR InterPro: IPR003663; Sugar_transporter.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGARTRANSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 2.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane.
 SQ SEQUENCE 489 AA; 53026 MW; D26436DAD69723C6 CRC64;

Query Match 33.4%; Score 777; DB 5; Length 489;
 Best Local Similarity 37.2%; Pred. No. 1,1e-36;
 Matches 172; Conservative 76; Mismatches 182; Indels 32; Gaps 6;

5 LAAPAAALGPLSGFALGYSSPAIPSLORAAPAPRLDDAASWFGAVVTGGAAGGV 64
 30 LAALSVSLGSLVGVFSATVMTDRNITSFEVTDAGSFWGICIMPLAGIAGG 89
 65 GMLVDNAGRKLSLLCSVPFAVAVTAADVWMLLGRLLTGLACGVALAPVYIE 124
 90 GPLEIYGLGRNTLLATPAVPIVSSLLIACAVNANAVLVCGRFLAGFCVGASLASLPYLDL 149
 125 IAVPAVAGLSCVOLAMVVGILLAYLAGWLEEMWMLAVLGVPSMLMLMCFEPETPR 184
 150 TVQPEVRAGTGLPTPAGNIGILLCTVAGSFMNMSHLATFGAALPPPELIMLLEPTEPR 209
 185 FLTLQHRROEA-TALRFLMGSEOGWE-----DPIGABOSFHALLRPGIYKP 232
 210 WFGVGRGLEERAKALKWLRGKEADVPEPLKGLMRSGQADDRQSRNTMELTLKLNUL-KP 268
 233 FIIGVSLMAFOQLSGVNAWVFATFEAKFK-KKSSLASVVGVIQVLEFTAALIMDR 291

[illegible]

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 Khan S., Kim C., Altati H., Bei B., Chin C., Chou J., Choi E.,
 Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
 Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharisky N.,
 Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 Theologis A., Ecker J.;
 RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC - SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL: AC025808; AAF79445.1;
 DR InterPro: IPR003662; sub_transporter.
 DR InterPro: IPR003663; Sugar_transporter.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGTRNSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 DR Transmembrane.
 KW S0
 SEQUENCE 515 AA: 55930 MW: 8057448064EPBD4 CRC64;

Query Match 31.4%; Score 730; DB 10; Length 515;
 Best local Similarity 36.3%; Pred. No. 5.3e-34;
 Matches 176; Conservative 81; Mismatches 172; Indels 56; Gaps 7;

QY 3 VFLLAAFAALGPISFGFALGYSSPAIPSLORAPAPRLDDAAASFGAVVTGAAAGV 62
 DB 49 VLACVLIVAGLPQIFQFTGCGYSPT---QAATTKDLGLTVSEYSGLSNNGAMGAI 104
 QY 63 LGGMVLDVDRAGKRLSLLLCSVPFVAGFAVITAADVMMLLGGRLTLGLACVAVSLAPVYI 122
 DB 105 ASQOIAEYGRKGSIMIAIPNIIGWLISFAKDTSEFLYGRLLBEGVGIIISYTPVYI 164
 QY 123 SEIAYAVANGLLGSCVQVLMVWVIGILAYLAGWLEMRMLAVLGCVPSPMLLMCEMPT 182
 DB 165 AETAPQNMKGALGSVNLSTVIGIMLAYLGLFVPMRIILAVLGLPCTLLIPGLFFIPES 224
 QY 183 PRELLTQHRR---QEAIALRFIMGSEQGE----- 209
 DB 225 PRMLVCLYNRIAYIVEFVALAKMGLTDFETSLQVLRGFETDITVEVNEIKVYTKLKC 284
 QY 210 -DEPICA---EFSFHLALLRQPIYKPFILIGVSLMAFOOLSGVNAWFAETIFEEAKK 265
 DB 285 FDSVSVSSSKRSVRFVDLAKRRRYRFPFLMWIGLALQOLGGINGVLFSTTFESAGVT 344
 QY 266 DSSLASVAVGVIOVLTAAVALIMDRAGRLILVLSGVWVFSTSAFAGFKLTGGPGN 325
 DB 345 SSVAVATFPGVGVVATGATMTLVKAGRRLLMISIGMTISLVIVAAAFYLKEVSPD 404
 QY 326 SSIVAIASAPVSAOPVDASVGLANLAVGSMCLPTAGFAVNGPITPMLMSEIFPLHKGVA 385
 DB 405 SNMYNIIISWVS-----VVGVAAMVIS--CSLGMGPIPLMIMSEILPVNIKGLA 450
 QY 386 TGCICVLTNNMLMAFLTKKSSLMELVRPYGAFPLASAFCLFSLVTLFCVPEPKGTLBQ 445
 DB 451 GSTATLTLNMFVSLVMTTAMMLL-AWSSSGTFTLVALVCGFTVAVSLWVPEPKGTLEE 509
 QY 446 ITAHF 450
 DB 510 IQALF 514

RESULT 13
 ID Q9FRL3
 AC Q9FRL3
 DT 01-MAR-2001 (TREMUREL. 16, Created)
 DT 01-MAR-2001 (TREMUREL. 16, Last sequence update)
 DE 01-JUN-2001 (TREMUREL. 17, Last annotation update)
 DE INTEGRAL MEMBRANE PROTEIN, PUTATIVE, 33518-36712.
 GN F22H5.6.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 RX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA.
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
 Maitl R., Ronald C.M., Koo H., Fujii C.Y., Utterback T.R.,
 Barstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC - SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL: AC025814; AAG12689.1;
 DR InterPro: IPR003662; sub_transporter.
 DR InterPro: IPR003663; Sugar_transporter.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGTRNSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
 DR Transmembrane.
 KW S0
 SEQUENCE 487 AA: 52901 MW: 98604E3642914472 CRC64;

Query Match 31.2%; Score 725.5; DB 10; Length 487;
 Best local Similarity 37.5%; Pred. No. 9.1e-34;
 Matches 173; Conservative 77; Mismatches 176; Indels 35; Gaps 7;

QY 3 VFLLAAFAALGPISFGFALGYSSPAIPSLORAPAPRLDDAAASFGAVVTGAAAGV 62
 DB 48 VLACVLIVAGLPQIFQFTGCGYSPT---QAATTKDLGLTVSEYSGLSNNGAMGAI 103
 QY 63 LGGMVLDVDRAGKRLSLLLCSVPFVAGFAVITAADVMMLLGGRLTLGLACVAVSLAPVYI 122
 DB 104 ASQOIAEYGRKGSIMIAIPNIIGWLISFAKDTSEFLYGRLLBEGVGIIISYTPVYI 163
 QY 123 SEIAYAVANGLLGSCVQVLMVWVIGILAYLAGWLEMRMLAVLGCVPSPMLLMCEMPT 182
 DB 164 AETAPQNMKGALGSVNLSTVIGIMLAYLGLFVPMRIILAVLGLPCTLLIPGLFFIPES 223
 QY 183 PRELLTQHRR---QEAIALRFIMGSEQGE-----FVSPDS-----DMYSLSL 229
 DB 224 PRMLAMGMDTDEFETSLQVLRGFE---TDITVEVNEIKRSVASTKRNTRFVDLAKRRY 280
 QY 230 YKPFILIGVSLMAFOOLSGVNAWFAETIFEEAKFKDSSLASVAVGVIOVLTAAVALIM 289
 DB 281 YPLMWIGLALQOLGGINGVLFSTTFESAGVTSSMAATFEGALQIVATISTMLV 340
 QY 290 DRAGRLILVLSGVWVFSTSAFAGFKLTGGPGNSSIVAIASAPVSAOPVDASVGLANL 349
 DB 341 DRAGRLILITSSVGTITSLVIVAAAFYLKE-----FVSPDS-----DMYSLSL 386
 QY 350 AVGSMCLPTAGFAVNGPITPMLMSEIFPLHKGVAATGICVLTNNMLMAFLTKKSSLM 409
 DB 387 SVYGVAAWVFSLGMPILPMLMSEILPVNIKGLAGSTATLANMFPSWLTMT-ANILL 445
 QY 410 VLRYGAEWLASAFCLFSLVTLFCVPEPKGTLDOITAHF 450
 DB 446 AWSSSGTFTLVGIVCAFTVAVTLMWPEPKGTLEEISL 486

RESULT 14
 Q9V848

ID Q9V848 PRELIMINARY; PRT: 465 AA.
AC Q9V848;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE CG6484 PROTEIN.
GN CG6484.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
CC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abil J.F., Abhyani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.A., Bouck J., Brooksstein P., Brothier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
RA Jajani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matile B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Wolley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL: AE003802; AAF57829.1; -;
DR FlyBase: FBgn0034247; CG6484.
DR InterPro: IPR003662; sub_transporter.
DR InterPro: IPR003663; Sugar transporter.
DR Pfam: PF00083; sugar_tr; 1
DR PRINTS: PR00171; SUGRTNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane.
SQ SEQUENCE 465 AA; 50271 MW; 004BACAF406BDBB CRC64;

Query Match 30.9%; Score 717.5; DB 5; Length 465;
Best Local Similarity 34.9%; Pred. No. 2.5e-33;
Matches 159; Conservative 95; Mismatches 172; Indels 25; Gaps 4;
4 FLAAFAAALCPISFGALGYSSPAIPSLQRAAPAPRLDDAAASWGAAYVTTLGAAGCYL 63

Db 10 YIAGLSAFCALCMGASIGWSSPVENMTVTYEXGFPRISSQFGWVSSLTLGAVTICP 69
Qy 64 GGMVLVDAGRKLSLLCSVPVAGFAVITAADYVMMGLLGLAGVASIAPYIS 123
Db 70 IGFADIMIGRPPTMLALPPYMGVLMLEPAKNVTMLFGFRFLMGCGACVAPAPMCT 129
Qy 124 EIVAPARGLGSCVQLMNVVGLLAVLAGVLEWRMLAVGCVPSLMLLMCFMPTP 183
Db 130 EITATALRGITGSGFOLLIVSGVLGYLVGAFLLPLTINICAILPYFAIHFMPEPSP 189
Qy 184 RELLQHRROPAI-ALPFLMSEQMED-----PPIGADSPHLALLRPGITKP 232
Db 190 VYLAMKGRNDPAKALQWLKRDADIDDELKEILEESQKQIDMQVNLSSLRPIVLKG 249
Qy 233 FIIGVSLMAFQOLSGVNAVMVYAEITFEENKFKDS-SLASVVGVIQVLAFAVALIMDR 291
Db 250 LGIAVLQVFOQWGTINAVLYSASIFEDTGSIDSGDATLIGTOVTSFLVAVALIDK 309
Qy 292 AGRLRLVLGCVWVFSTSAFGAYFKLTGGPGNSHVASIPYSAOPVDASVGLAVLAV 351
Db 310 AGRRILLISGVLMVAVSTALMGVFEQLENDPASMDN-----FGWLPI 352
Qy 352 GSMCLFIAGVAGCPLIPMLMSEIFPLHYKGVATGICVLTNMLAPLVTKFESSLMVL 411
Db 353 SSICIFITFPISIGPGPVWIMAELEFSEDKSVAGSIAGTSNMLSAFVNTLLPFLKSSI 412
Qy 412 RPYGAFMLASAFICFVSILFTFCFPETKGTLDEI 446
Db 413 GCPFTFWIFATAIVIAFFYSLEFVPETKGTTIEL 447
RESULT 15
Q39416
ID Q39416 PRELIMINARY; PRT: 490 AA.
AC Q39416;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE INTEGRAL MEMBRANE PROTEIN.
OS Beta vulgaris (Sugar beet).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllales; Chenopodiaceae; Beta.
OX NCBI_TaxID=3555;
RN [1]
RP SEQUENCE FROM N.A.
RA Chlou T.J., Bush D.R.;
RL Plant Physiol. 110:0-0(0).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL: U43629; AAB53155.1; -;
DR Mendel: 16093; Belvu:2447;16093.
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN.1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transporter; Transmembrane.
SQ SEQUENCE 490 AA; 53020 MW; 3F8C96D26989CBBB CRC64;

Query Match 30.9%; Score 717.5; DB 10; Length 490;
Best Local Similarity 36.1%; Pred. No. 2.6e-33;
Matches 167; Conservative 82; Mismatches 175; Indels 39; Gaps 6;
Qy 3 VFLAFAAALCPISFGALGYSSPAIPSLQRAAPAPRLDDAAASWGAAYVTTLGAAGCY 62
Db 50 VLACVLIVAGLPQIGFTAGVSPT---QSAITNELGLSVAEYSWGCSLSNVAAMGAI 105
Qy 63 LGGWLVDRARCKLSLLCSVPVAGFAVITAADYVMMGLLGLAGVASIAPYI 122
Db 106 ASQDISFYIKRGSMLIAIPNITIGMLAISFAKSSFLYMKRMLEGRGVGIISTVYPI 165

Thu Feb 14 07:44:56 2002

us-09-516-493-7.rspt

Page 10

```

OY 123 SEIAPVARGLLSGCVOLMVAVGLILVAGVLEMMLEMLVGCVPSEMLLMCMCEPBT 182
    | | | : : : | | | : : : | | | : : : | | | : : : | | | : : :
Db 166 SEIAPONIRGALGAVGNOLSVTIGIMLSVMLOGLFVPHILILAVGLPCTILLIPGLFIPE 225
    | | | : : : | | | : : : | | | : : : | | | : : : | | | : : :
OY 183 PRELLTQHRBOEAIALRFLMGSQDCMEDPPICAE-----QSFHALLROP 227
    | | | : : : | | | : : : | | | : : : | | | : : : | | | : : :
Db 226 PRMLAKMGMEEFEFSLQVLNKG-----DDDISLEVNEIKRSVAASSKRTTIRPAELROR 280
    | | | : : : | | | : : : | | | : : : | | | : : : | | | : : :
OY 228 GIYKFFILGVSLMAFQOOLSGVNAVMFYAEITFEFEKFRDSSLASVVGVIOVLEPTAAAL 287
    | | | : : : | | | : : : | | | : : : | | | : : : | | | : : :
Db 281 RYMLPLMIGNGLLILQOLSGINGVLFPYSTIFKEAGVTSMAATFGAGAOVIATVYTTW 340
    | | | : : : | | | : : : | | | : : : | | | : : : | | | : : :
OY 288 IMDRAGRLLILVSGVAVVFTSAFCAVFKILQCGPQNSSHVALSAPVSAQPDASVGLA 347
    | | | : : : | | | : : : | | | : : : | | | : : : | | | : : :
Db 341 LVDSGRRLLLIVSSGTTLSLIVVAMSFLEKENVSDSESTVSY-----FS 386
    | | | : : : | | | : : : | | | : : : | | | : : : | | | : : :
OY 348 WLAVGSMCLFIAGCAVNGCPIPRMLMSEIEPLHAKGVAATGICVTLMLMAFLVTKRESSL 407
    | | | : : : | | | : : : | | | : : : | | | : : : | | | : : :
Db 387 ILSVGVAAVAVVFTSLGIGALPWIITSEILPINIKGLAGSIATLANFVAMVITMTANI 445
    | | | : : : | | | : : : | | | : : : | | | : : : | | | : : :
OY 408 MEVLPRYCAFMLASAFCSFSLVFLFCVPEPKKCTEQQITAHF 450
    | | | : : : | | | : : : | | | : : : | | | : : : | | | : : :
Db 446 MLSMNSGGTFSIYMWCAFYVAFAVIVAVPEPKGTTELIQMSF 488
    | | | : : : | | | : : : | | | : : : | | | : : : | | | : : :

```

Search completed: February 13, 2002, 22:04:42
Job time: 863 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 21:50:11 : Search time 104.41 Seconds

(without alignments)
117.059 Million cell updates/sec

Title: US-09-516-493-10

Perfect score: 889

Sequence: 1 WRPLZNMIPRAGIRHMLV.....VTGSPELISWLRVSKGVEMPO 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /SID52/gcgdata/geneseq/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	627	70.5	478	22	AA66933
2	598	67.3	478	22	AA66939
3	590	66.4	477	22	AA66934
4	549	61.8	477	22	AA66937
5	540	60.7	477	22	AA66932
6	281	31.6	507	22	AA66937
7	279.5	31.4	507	22	AA66941
8	278	31.3	105	21	AA95019
9	278	31.3	106	21	AA645158
10	273.5	31.0	147	21	AA95032
11	267	30.0	503	22	AA66938

12	257	28.9	193	21	AA637400	Arabidopsis thalia
13	257	28.9	193	21	AA637422	Arabidopsis thalia
14	257	28.9	212	21	AA637398	Arabidopsis thalia
15	257	28.9	212	21	AA637420	Arabidopsis thalia
16	257	28.9	437	21	AA637350	Arabidopsis thalia
17	257	28.9	437	21	AA620978	Arabidopsis thalia
18	257	28.9	438	21	AA638871	Arabidopsis thalia
19	257	28.9	447	21	AA63749	Arabidopsis thalia
20	257	28.9	447	21	AA620977	Arabidopsis thalia
21	257	28.9	448	21	AA638870	Arabidopsis thalia
22	257	28.9	463	21	AA633748	Arabidopsis thalia
23	257	28.9	463	21	AA620976	Arabidopsis thalia
24	257	28.9	464	21	AA638869	Arabidopsis thalia
25	235.5	26.5	526	20	AA602168	A facilitative glu
26	212.5	23.9	471	21	AA648646	Arabidopsis thalia
27	212.5	23.9	558	21	AA648645	Arabidopsis thalia
28	212.5	23.9	564	21	AA648644	Arabidopsis thalia
29	209.5	23.6	420	21	AA630434	Arabidopsis thalia
30	209.5	23.6	477	21	AA630433	Arabidopsis thalia
31	209.5	23.6	483	21	AA630432	Arabidopsis thalia
32	209	23.5	369	21	AA610667	Arabidopsis thalia
33	209	23.5	375	21	AA610666	Arabidopsis thalia
34	209	23.5	382	21	AA610665	Arabidopsis thalia
35	209	23.5	398	21	AA639283	Arabidopsis thalia
36	209	23.5	468	21	AA639282	Arabidopsis thalia
37	209	23.5	470	21	AA639281	Arabidopsis thalia
38	198.5	22.3	237	21	AA606143	Arabidopsis thalia
39	198.5	22.3	294	21	AA606142	Arabidopsis thalia
40	198.5	22.3	362	21	AA606141	Arabidopsis thalia
41	198.5	22.3	473	21	AA652317	Arabidopsis thalia
42	198.5	22.3	491	21	AA652316	Arabidopsis thalia
43	198.5	22.3	522	21	AA652315	Arabidopsis thalia
44	197.5	22.2	473	21	AA616800	Arabidopsis thalia
45	197.5	22.2	491	21	AA616799	Arabidopsis thalia

ALIGNMENTS

RESULT 1	AA66933	standard; Protein: 478 AA.
ID	AA66933	
AC	AA66933	
XX		
DT	17-APR-2001	(first entry)
XX		
DE	Rat GLUTX1.	
XX		
KW	Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;	
KW	hexose transport disorder; ischaemia; diabetes; hyperglycaemia;	
KW	hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.	
XX		
OS	Rattus sp.	
XX		
PN	W0200104145-A2.	
XX		
PD	18-JAN-2001.	
XX		
PE	14-JUL-2000; 2000WO-IB01042.	
XX		
PR	14-JUL-1999; 99US-0143907.	
PR	27-AUG-1999; 99US-0151140.	
PR	23-FEB-2000; 2000US-0184285.	
PR	13-JUL-2000; 2000US-0616132.	
XX		
PA	(UCLA-) UNIV LAUSANNE.	
XX		
PI	Thorens B, Ibberson M, Uldry M;	
XX	WPI; 2001-112615/12.	
DR	N-PSDB; AAF55866.	
XX		

PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -
 XX
 PS Claim 11: Page 71-72; 124pp; English.
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is rat GLUTX1.
 CC
 XX
 SQ Sequence 478 AA:
 Query Match 70.5%; Score 627; DB 22; Length 478;
 Best Local Similarity 100.0%; Pred. No. 3,1e-64;
 Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 19 LVPISAPADVHLGLAMLAAGSMCLFIAGFAGVGMGPIMLIMSEITPPLHKGATGCVL 78
 Db 357 lvpisapadvhlglamlavagsmclfiagfavgwgpipwlmselfplhkvatgvcyl 416
 QY 79 TNNFMAFLVTKFENSIMELIRPYGAFMLTAFCILSVLFTLPVPTKGRTEQITAH 136
 Db 417 tnnfmafvltkfensimeilrpygafwlttafcilsvlftltpvptkgrtleqitah 474
 RESULT 2
 AAB66939
 ID AAB66939 standard; Protein; 478 AA.
 XX
 AC AAB66939;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE GLUTX1 consensus sequence.
 XX
 KM GLUTX; gene therapy; vaccine; hexose transport modulator; human; rat;
 KM hexose transport disorder; ischaemia; diabetes; hyperglycaemia; murine;
 KM hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 OS Homo sapiens.
 OS Mus sp.
 OS Rattus sp.
 XX
 PN WO200104145-A2.
 XX
 PD 18-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-IB01042.
 XX
 PR 14-JUL-1999; 99US-0143907.
 PR 27-AUG-1999; 99US-0151140.
 PR 23-FEB-2000; 2000US-0184285.
 PR 13-JUL-2000; 2000US-0616132.
 XX
 PA (UYLA-) UNIV LAUSANNE.
 XX
 PI Thorens B, Ibbersson M, Uldry M;
 XX
 DR WPI: 2001-112615/12.
 XX
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -
 XX
 PS Claim 11: Page 74-75; 124pp; English.
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative

CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is a consensus sequence
 CC for GLUTX1.
 CC
 XX
 SQ Sequence 478 AA:
 Query Match 67.3%; Score 598; DB 22; Length 478;
 Best Local Similarity 93.2%; Pred. No. 7,2e-61;
 Matches 110; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 QY 19 LVPISAPADVHLGLAMLAAGSMCLFIAGFAGVGMGPIMLIMSEITPPLHKGATGCVL 78
 Db 357 lvpisapadvhlglamlavagsmclfiagfavgwgpipwlmselfplhkvatgvcyl 416
 QY 79 TNNFMAFLVTKFENSIMELIRPYGAFMLTAFCILSVLFTLPVPTKGRTEQITAH 136
 Db 417 tnnfmafvltkfensimeilrpygafwlttafcilsvlftltpvptkgrtleqitah 474
 RESULT 3
 AAB66934
 ID AAB66934 standard; Protein; 477 AA.
 XX
 AC AAB66934;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE Murine GLUTX1.
 XX
 KM Murine; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KM hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
 KM hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 OS Mus sp.
 OS
 PN WO200104145-A2.
 XX
 PD 18-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-IB01042.
 XX
 PR 14-JUL-1999; 99US-0143907.
 PR 27-AUG-1999; 99US-0151140.
 PR 23-FEB-2000; 2000US-0184285.
 PR 13-JUL-2000; 2000US-0616132.
 XX
 PA (UYLA-) UNIV LAUSANNE.
 XX
 PI Thorens B, Ibbersson M, Uldry M;
 XX
 DR WPI: 2001-112615/12.
 DR N-PSDB: AAF55867.
 XX
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -
 XX
 PS Claim 11: Page 73-74; 124pp; English.
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is murine GLUTX1.
 CC
 XX
 SQ Sequence 477 AA:

Query Match	Score	DB	Length
Query Match	66.4%	Score 590	DB 22
Best Local Similarity	90.7%	Pred. No. 6	1e-60
Matches 107	Conservative	7	Mismatches 4
			Indels 0
			Gaps 0
OY	19	LVPISAEPPDHYHLGAMILAVGSMCFLFIACFAVGMGIPMLISEIPPLTHKGAATGCVAL	78
DB	356	LYPIAEPDYGVYGLAWLAVGSMCLFIAGFAVGGIPWLLMSEIPHLVKKYATGICVAL	415
OY	79	TNNMFAPLVTKEFNSIMELIRPYGAFWMLTAACILISVLEPTITFPVETKSGTYLEQITAH	136
DB	416	LNWFMAFLVTKEFNSYVMEMLIRPYGAFWMLTAACILISVLEPTITFPVETKSGTYLEQITAH	473
RESULT	4		
ID	AAE04888	standard; Protein; 477 AA.	
AC	AAE04888		
XX	10-SEP-2001	(First entry)	
DE	Human transporter and ion channel-1 (TRICH-1) protein.		
XX			
KW	Human; transporter and ion channel-1; TRICH-1; vaccine; cystic fibrosis;		
KW	gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy;		
KW	hyperextension; angina; neurological disorder; asthma; bipolar disorder;		
KW	dementia; depression; Alzheimer's disease; epilepsy; mood; arhythmia;		
KW	Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke;		
KW	Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy;		
KW	demyelinating disease; mental disorder; Schizophrenia; polymyositis;		
KW	muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease;		
KW	dermatomyositis; diabetes mellitus; immunological disorder; psoriasis;		
KW	rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus;		
KW	sickle cell anaemia; Wilson's disease; infertility; Cushing's disease;		
KW	scroderma; pulmonary artery stenosis; noctropic; Addison's disease;		
KW	malabsorption syndrome; hypercholesterolaemia; cancer.		
XX			
OS	Homo sapiens.		
XX			
XX			
FT	Key	Location/Qualifiers	
FT	Domain	29..474	
FT		/note="Sugar transporter domain"	
FT	Domain	259..279	
FT		/label="Transmembrane_domain	
FT	Domain	293..313	
FT		/label="Transmembrane_domain	
FT	Domain	320..339	
FT		/label="Transmembrane_domain	
FT	Domain	438..457	
FT		/label="Transmembrane_domain	
XX			
XX	WO200146258-A2.		
PN			
PD	28-JUN-2001.		
XX			
XX	22-DEC-2000; 2000WO-US35095.		
PE			
XX	23-DEC-1999; 99US-0172000.		
PR	14-JAN-2000; 2000US-0176083.		
PR	21-JAN-2000; 2000US-0177332.		
PR	28-JAN-2000; 2000US-0178572.		
PR	02-FEB-2000; 2000US-0179756.		
PR	10-FEB-2000; 2000US-0181625.		
XX			
PA	(INCYT) INCYTE GENOMICS INC.		
PI	Baughn MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy R, Lal P;		
PI	Hillman JL, Azinzaiz Y, Yue H, Nguyen DB, Yao MG, Gandhi AR;		
PI	Tang YI, Khan FA;		
XX			
XX	WPI; 2001-418042/4.		
DR	N-PSDB; AAD09552.		

[illegible]

PA (UYLA-) UNIV LAUSANNE.
 XX
 XX Thorens B, Ibberson M, Uldry M;
 XX
 DR MPI: 2001-112615/12.
 DR N-PSDB: AAF55865.
 XX
 XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -
 XX
 PS Claim 11: Page 70-71; 124pp; English.
 XX
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is human GLUTX1.
 CC
 XX
 SQ Sequence 477 AA:
 Query Match 60.7%; Score 540; DB 22; Length 477;
 Best Local Similarity 82.8%; Pred. No. 3.9e-54;
 Matches 96; Conservative 11; Mismatches 9; Indels 0; Gaps 0;
 QY 21 PISAEPADVHLGLAVLAVGSMCLFTAGFVAGMGPIPLMLSEIFPLHICVATGCVLTN 80
 Db 358 pvsagvvdasvsglawlavagmclftagfvgwgpilmlseilplhvkvavaglcvtln 417
 QY 81 WFMAPLVTEKFNSEIMELRPYGAFWLTAFCILSVFTLFEVPEYGRTEQITAH 136
 Db 418 wlmeflvtkfesslmelvrlpygafwlasafclsvlftlcypelngktleqltn 473
 RESULT 6
 AAB66937
 ID AAB66937 standard; Protein; 507 AA.
 XX
 AC AAB66937;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE Human GLUTX3.
 XX
 KW Human; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200104145-A2.
 XX
 PD 18-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-1B01042.
 XX
 PR 14-JUL-1999; 99US-0143907.
 PR 27-AUG-1999; 99US-0151140.
 PR 23-FEB-2000; 2000US-0184285.
 PR 13-JUL-2000; 2000US-0616132.
 XX
 PA (UYLA-) UNIV LAUSANNE.
 XX
 XX Thorens B, Ibberson M, Uldry M;
 XX
 DR MPI: 2001-112615/12.
 DR N-PSDB: AAF55870.
 XX
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,

PT e.g. ischemia and diabetes -
 XX
 XX Claim 11; Page 81-82; 124pp; English.
 XX
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is human GLUTX3.
 CC
 XX
 SQ Sequence 507 AA:
 Query Match 31.6%; Score 281; DB 22; Length 507;
 Best Local Similarity 44.3%; Pred. No. 4.7e-24;
 Matches 62; Conservative 21; Mismatches 33; Indels 4; Gaps 2;
 QY 2 RPLZNZMIPRAGIRHETLPISEPADVHLG-LAVLAVGSMCLFTAGFVAGMGPIPLML 60
 Db 365 rplspn---stagleseswgdlaqplaapayltlvpllatmlflmgvavgwptlilm 421
 QY 61 SEIFPLHIGVATGVCVLTNMFMAFLVTKFNSEIMELRPYGAFWLTAFCILSVFTLT 120
 Db 422 sevlplrtargvasgclvaswtatfvltpvsvftglqvpflffaalcivslvftgc 481
 QY 121 FVPEYGRTEQITAHLRDG 140
 Db 482 cvpetkgrslqeslftfyg 501
 RESULT 7
 AAB66941
 ID AAB66941 standard; Protein; 507 AA.
 XX
 AC AAB66941;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE GLUTX3 consensus sequence.
 XX
 KW GLUTX; gene therapy; vaccine; hexose transport modulator; human; rat;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX
 OS Homo sapiens.
 OS Rattus sp.
 XX
 PN WO200104145-A2.
 XX
 PD 18-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-1B01042.
 XX
 PR 14-JUL-1999; 99US-0143907.
 PR 27-AUG-1999; 99US-0151140.
 PR 23-FEB-2000; 2000US-0184285.
 PR 13-JUL-2000; 2000US-0616132.
 XX
 PA (UYLA-) UNIV LAUSANNE.
 XX
 XX Thorens B, Ibberson M, Uldry M;
 XX
 DR MPI: 2001-112615/12.
 XX
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -
 XX
 PS Claim 11; Page 83-84; 124pp; English.
 XX
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and

CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is a consensus sequence
 CC for GLUTX3.

CC Sequence 507 AA:

Query Match 31.4%; Score 279.5; DB 22; Length 507;
 Best Local Similarity 45.7%; Pred. No. 7.1e-24;
 Matches 59; Conservative 21; Mismatches 48; Indels 1; Gaps 1;

QY 13 AGIRHETLPSAEPADVHIG-LAMLAVGSNCFTIAGPAVGWGPFPMLMSEIFPLHKGV 71
 DB 373 agleseswgdlaqpaaapagyltlvpllatmflimgavqwgpidwlmsevipfargy 432
 QY 72 ATGVCLTMMFMAFLVTKFENSIMETLRPYGAFWLTAFCLISLFTLTFVPEPKGRTLE 131
 DB 433 asglcvlaswltatvltksflpvsstfglyqvpflffaalcslvstfgccvpetkgrsle 492
 QY 132 QTTAHLRDG 140
 DB 493 qiesffrtg 501

RESULT 8

AA95019
 ID AAY95019 standard; Protein: 105 AA.

XX AAY95019;
 DT 19-JUN-2000 (first entry)

DE Human secreted protein vq1_1, SEQ ID NO:78.

KW Human; secreted protein; cancer; tumour; cardiovascular disorder;
 KW blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;
 KW infection; fungal; bacterial; viral; HIV; allergy; arthritis;
 KW neurodegenerative disease; asthma; contraceptive.

OS Homo sapiens.

PN WO200011015-A1.

XX 02-MAR-2000.

XX 24-AUG-1999; 99WO-US19351.

XX 24-AUG-1998; 98US-0097638.
 PR 24-AUG-1998; 98US-0097659.
 PR 09-SEP-1998; 98US-0099618.
 PR 28-SEP-1998; 98US-0102092.
 PR 25-NOV-1998; 98US-0109978.
 PR 23-DEC-1998; 98US-0113645.
 PR 23-DEC-1998; 98US-0113646.
 PR 23-AUG-1999; 99US-0379246.

XX (ALPH-) ALPHAGENE INC.

XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;

XX WPI: 2000-224657/19.

XX New secreted or transmembrane proteins and polynucleotides encoding
 PT them, useful for treating neurodegenerative disorders, autoimmune
 PT diseases and cancer -

PS Claim 87; Page 335-336; 357pp; English.

XX The invention relates to 40 human secreted proteins (AAY94981-Y95020),

CC and cDNA sequences encoding them (AAA23423-A23462). The secreted
 CC proteins of the invention include those that are thought to be only
 CC partially secreted, i.e., transmembrane proteins. The proteins of the
 CC invention may exhibit one or more activities selected from the following:

CC cytokine activity; cell proliferation; differentiation; immune
 CC modulation; haematopoiesis regulation; tissue growth activity;
 CC activating/inhibiting activity; chemotactic/chemokinetic activity; haemostatic
 CC and thrombolytic activity; anti-inflammatory activity; and tumour
 CC inhibition activity. The proteins may be administered to patients as
 CC vaccines, and the nucleotides may be used as part of a gene therapy
 CC regime. Diseases or conditions that may be treated using the proteins or
 CC nucleotides of the invention include autoimmune diseases; genetic
 CC disorders; haemophilia; cardiovascular diseases; cancer; bacterial,
 CC fungal and viral infections, especially HIV; multiple sclerosis;
 CC rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome;
 CC insulin dependent diabetes mellitus; and allergic reactions such as
 CC asthma and anaemia. They may also be used for treating wounds, burns,
 CC ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease and amyotrophic
 CC lateral sclerosis (ALS). Proteins with activating/inhibiting activity may
 CC additionally be useful as contraceptives. Nucleic acid sequences of the
 CC invention may be used in chromosome mapping, and as a source of
 CC diagnostic primers and probes. The present sequence represents one of the
 CC 40 proteins of the invention.

CC Sequence 105 AA:

Query Match 31.3%; Score 278; DB 21; Length 105;
 Best Local Similarity 53.1%; Pred. No. 1.4e-24;
 Matches 52; Conservative 17; Mismatches 29; Indels 0; Gaps 0;

QY 43 LFIAGFAVGWGPFPMLMSEIFPLHKGVATGVCLTMMFMAFLVTKFENSIMETLRPYG 102
 DB 2 limgavqwgpidwlmsevlprargvasglcvlaswltatvltksflpvsstfglyq 61

QY 103 AFMLTAFCILSVLFTLTFVPEPKGRTLEQITAHLRDG 140
 DB 62 pfffaalcslvstfgccvpetkgrsleqiesffrtg 99

RESULT 9

ID AAB45158 standard; Protein: 106 AA.

XX AAB45158;

XX 12-FEB-2001 (first entry)

DE Human secreted protein sequence encoded by gene 39 SEQ ID NO:99.

KW Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;
 KW antiproliferative; cytostatic; cardiact; vasotropic; cerebroprotective;
 KW neutrotropic; neuroprotective; antibacterial; virucide; fungicide; cancer;
 KW ophthalmological; autoimmune disease; hyperproliferative disorder;
 KW cardiovascular disorder; cerebrovascular disorder; wound healing;
 KW nervous system disorder; aging; chemotaxis.

XX Homo sapiens.

XX WO200058467-A1.

XX 05-OCT-2000.

XX 22-MAR-2000; 2000WO-US07505.

XX 26-MAR-1999; 99US-0126502.
 PR 17-DEC-1999; 99US-0172410.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX

DR WPI: 2000-611712/58.
 DR N-PSDB; AAC80569.
 XX
 PT Nucleic acid molecules encoding human secreted proteins, used in
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 PT Parkinson's diseases and cancers -
 PS Claim 11: Page 396; 440pp; English.
 XX
 CC Polynucleotide sequences AAC80531-C80580 represent cDNA encoding human
 CC secreted proteins AAB45120-B45169. Sequences AAB45170-B45225 represent
 CC alternative polypeptides encoded by the genes, and amino acid sequences
 CC to which they are homologous. The genes and proteins have activities
 CC dependent on the tissues and cells in which they are expressed. Examples
 CC of their activities include immunosuppressive; antirheumatic;
 CC antineoplastic; antiproliferative; cytostatic; cardiac; vasotropic;
 CC cerebroprotective; neurotrophic; neuroprotective; antibacterial; virucide;
 CC fungicide; and ophthalmological. The secreted proteins, polynucleotides,
 CC antagonists and agonists may be useful in treating, preventing and/or
 CC diagnosing diseases and disorders such as autoimmune diseases
 CC e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms
 CC of the breast or liver, cardiovascular disorders e.g. cardiac arrest,
 CC cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
 CC The polypeptides can also be used to aid wound healing and epithelial
 CC cell proliferation, to prevent skin aging due to sunburn, to maintain
 CC organs before transplantation, for supporting cell culture of primary
 CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
 CC also be used as a food additive or preservative to increase or decrease
 CC storage capabilities. AAC80522-C80530 and AAB45119 represent sequences
 CC used in the isolation and characterisation of the genes and proteins of
 CC the invention.
 XX
 XX
 S0 Sequence 106 AA:
 31.3%; Score 278; DB 21; Length 106;
 Query Match Best Local Similarity 53.1%; Pred. No. 1.4e-24;
 Matches 52; Conservative 17; Mismatches 29; Indels 0; Gaps 0;
 QY 43 LFIAGFAVWGCPFWLMSSEIFPLHKGAVGCVLTNMFMAFLVTKFNSIMEILRPYG 102
 Db 2 Lfimgyavwgpltwlmssevlplrarvgasglcvlswlafatltksflpvsstfglqv 61
 QY 103 AFMLTAACILSVLFTLFPETKGRLEDTAHLRDG 140
 Db 62 pffifaalcivslvtgcvpetkyslqslqefitrg 99
 RESULT 10
 AAY95032
 ID AAY95032 standard; Protein: 147 AA.
 XX
 AC AAY95032;
 XX
 DT 19-JUN-2000 (first entry)
 XX
 DE Human clone vql_1 insertional variant ORF, SEQ ID NO:132.
 XX
 KW Human: secreted protein; cancer; tumour; cardiovascular disorder;
 KW blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;
 KW infection; fungal; bacterial; viral; HIV; allergy; arthritis;
 KW neurodegenerative disease; asthma; contraceptive; open reading frame;
 KM ORF; variant.
 XX
 OS Homo sapiens.
 XX
 PN WO200011015-A1.
 XX
 PD 02-MAR-2000.
 XX
 PF 24-AUG-1999; 99WO-USI9351.

XX
 PR 24-AUG-1998; 98US-0097638.
 PR 24-AUG-1998; 98US-0097659.
 PR 09-SEP-1998; 98US-0099618.
 PR 28-SEP-1998; 98US-0102092.
 PR 25-NOV-1998; 98US-0109978.
 PR 23-DEC-1998; 98US-0113645.
 PR 23-DEC-1998; 98US-0113646.
 PR 23-AUG-1999; 99US-0379246.
 PA (ALPH-) ALPHAGENE INC.
 PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
 PT WPI: 2000-224657/19.
 PT
 PT New secreted or transmembrane proteins and polynucleotides encoding
 PT them, useful for treating neurodegenerative disorders, autoimmune
 PT diseases and cancer -
 PS Disclosure: Page 351-352; 357pp; English.
 XX
 CC The invention relates to 40 human secreted proteins (AAY94981-Y95020),
 CC and cDNA sequences encoding them (AAB23423-A23462). The secreted
 CC proteins of the invention include those that are thought to be only
 CC partially secreted, i.e., transmembrane proteins. The proteins of the
 CC invention may exhibit one or more activities selected from the following:
 CC cytokine activity; cell proliferation; differentiation; immune
 CC modulation; haematopoiesis regulation; tissue growth activity;
 CC activin/inhibin activity; chemotactic/chemokinetic activity; haemostatic
 CC and thrombolytic activity; anti-inflammatory activity; and tumour
 CC inhibition activity. The proteins may be administered to patients as
 CC vaccines, and the nucleotides may be used as part of a gene therapy
 CC regime. Diseases or conditions that may be treated using the proteins or
 CC nucleotides of the invention include autoimmune diseases; genetic
 CC disorders; haemophilia; cardiovascular diseases; cancer; bacterial,
 CC fungal and viral infections, especially HIV; multiple sclerosis;
 CC rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome;
 CC insulin dependent diabetes mellitus; and allergic reactions such as
 CC asthma and anaemia. They may also be used for treating wounds, burns,
 CC ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease and amyotrophic
 CC lateral sclerosis (ALS). Proteins with activin/inhibin activity may
 CC additionally be useful as contraceptives. Nucleic acid sequences of the
 CC invention may be used in chromosome mapping, and as a source of
 CC diagnostic primers and probes. Sequences AAY95024-Y95025, AAY95029 and
 CC AAY95031-Y95032 represent additional open reading frames (ORFs) that are
 CC encoded by deletional or insertional variants of the cDNA clones of the
 CC invention.
 XX
 XX
 S0 Sequence 147 AA:
 31.0%; Score 275.5; DB 21; Length 147;
 Query Match Best Local Similarity 50.5%; Pred. No. 4.1e-24;
 Matches 56; Conservative 17; Mismatches 37; Indels 1; Gaps 1;
 QY 43 LFIAGFAVWGCPFWLMSSEIFPLHKGAVGCVLTNMFMAFLVTKFNSIMEILRPYG 102
 Db 2 Lfimgyavwgpltwlmssevlplrarvgasglcvlswlafatltksflpvsstfglqv 61
 QY 103 AFMLTAACILSVLFTLFPETKGRLEDTAHLRDGSPLSVTG-SPFL 152
 Db 62 pffifaalcivslvtgcvpetkyslqslqefitrgpgekvllalggppl 112
 RESULT 11
 AAB66938
 ID AAB66938 standard; Protein: 503 AA.
 XX
 AC AAB66938;
 XX
 DT 17-APR-2001 (first entry)

XX Rat GLUTX3.
DE
XX
KW Rat: GLUTX: gene therapy: vaccine: hexose transport modulator;
KM hexose transport disorder: ischaemia: diabetes: hyperglycaemia;
XX hypoglycaemia: glucose metabolism disorder: neurodegenerative disease.
OS Rattus sp.
XX
XX MO200104145-A2.
PN
XX
XX 18-JAN-2001.
PD
XX
XX 14-JUL-2000; 2000WO-1B01042.
PF
XX
XX 14-JUL-1999; 99US-0143907.
PR
XX 27-AUG-1999; 99US-0151140.
PR
XX 23-FEB-2000; 2000US-0184285.
PR
XX 13-JUL-2000; 2000US-0616132.
XX
XX (UYLA-) UNIV LAUSANNE.
PA
XX
XX Thorens B, Ibberson M, Uldry M;
PI
XX WPI: 2001-112615/12.
DR
XX N-PSDB; AAF55871.
XX
XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
PT e.g. Ischemia and diabetes -
XX
XX Claim 11: Page 82-83; 124pp: English.
XX
XX The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischaemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is rat GLUTX3.
XX
XX
XX Sequence 503 AA:
SO

Query Match 30.0%; Score 267; DB 22; Length 503;
Best Local Similarity 42.7%; Pred. No. 2e-22;
Matches 61: Conservative 22; Mismatches 42; Indels 18; Gaps 3;

QY 14 GIRHELVPISAEPAD-----VHLG-----LAWLAVGSMCLFIAGFAVGWGPWPW 57
DB 355 glyvqlyprltpnsteiveltgqtpaaafnylltllpilatmlfmygmwqplw 414
QY 58 LMSRIFPLHIKGVATGCVLTNMFMAFLVTKFENSIMELKPYCAFMTAAFCILSVLF 117
DB 415 lhmsevlplrarvgasglcvlswlatvltkylilavnaiglyvplftsaicllslf 474
QY 118 TLTFVETKGRLEQITFA--HLR 138
DB 475 tgcvcyptcgrslqleaftfhr 497

RESULT 12
AAG37400
ID AAG37400 standard; Protein: 193 AA.
XX
XX AAG37400;
AC
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 45980.
DE
XX Protein identification: signal transduction pathway: metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
PR

KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR
XX 05-MAR-1999; 99US-0123180.
PR
XX 09-MAR-1999; 99US-0123548.
PR
XX 23-MAR-1999; 99US-0125788.
PR
XX 25-MAR-1999; 99US-0126264.
PR
XX 29-MAR-1999; 99US-0126785.
PR
XX 01-APR-1999; 99US-0127462.
PR
XX 06-APR-1999; 99US-0128234.
PR
XX 08-APR-1999; 99US-0128714.
PR
XX 16-APR-1999; 99US-0129845.
PR
XX 19-APR-1999; 99US-0130077.
PR
XX 21-APR-1999; 99US-0130449.
PR
XX 23-APR-1999; 99US-0130510.
PR
XX 23-APR-1999; 99US-0130891.
PR
XX 28-APR-1999; 99US-0131449.
PR
XX 30-APR-1999; 99US-0132048.
PR
XX 30-APR-1999; 99US-0132407.
PR
XX 04-MAY-1999; 99US-0132484.
PR
XX 05-MAY-1999; 99US-0132485.
PR
XX 06-MAY-1999; 99US-0132486.
PR
XX 06-MAY-1999; 99US-0132487.
PR
XX 07-MAY-1999; 99US-0132863.
PR
XX 11-MAY-1999; 99US-0134256.
PR
XX 14-MAY-1999; 99US-0134219.
PR
XX 14-MAY-1999; 99US-0134221.
PR
XX 14-MAY-1999; 99US-0134370.
PR
XX 18-MAY-1999; 99US-0134768.
PR
XX 19-MAY-1999; 99US-0134941.
PR
XX 20-MAY-1999; 99US-0135124.
PR
XX 21-MAY-1999; 99US-0135153.
PR
XX 24-MAY-1999; 99US-0135629.
PR
XX 25-MAY-1999; 99US-0136021.
PR
XX 27-MAY-1999; 99US-0136392.
PR
XX 28-MAY-1999; 99US-0136782.
PR
XX 01-JUN-1999; 99US-0137222.
PR
XX 03-JUN-1999; 99US-0137528.
PR
XX 04-JUN-1999; 99US-0137502.
PR
XX 07-JUN-1999; 99US-0137724.
PR
XX 08-JUN-1999; 99US-0138094.
PR
XX 10-JUN-1999; 99US-0138540.
PR
XX 10-JUN-1999; 99US-0138847.
PR
XX 14-JUN-1999; 99US-0139119.
PR
XX 16-JUN-1999; 99US-0139452.
PR
XX 16-JUN-1999; 99US-0139453.
PR
XX 17-JUN-1999; 99US-0139492.
PR
XX 18-JUN-1999; 99US-0139454.
PR
XX 18-JUN-1999; 99US-0139455.
PR
XX 18-JUN-1999; 99US-0139456.
PR
XX 18-JUN-1999; 99US-0139457.
PR
XX 18-JUN-1999; 99US-0139458.
PR
XX 18-JUN-1999; 99US-0139459.
PR
XX 18-JUN-1999; 99US-0139460.
PR
XX 18-JUN-1999; 99US-0139461.
PR
XX 18-JUN-1999; 99US-0139462.
PR
XX 18-JUN-1999; 99US-0139463.
PR
XX 18-JUN-1999; 99US-0139750.
PR
XX 18-JUN-1999; 99US-0139763.
PR
XX 21-JUN-1999; 99US-0139817.
PR
XX 22-JUN-1999; 99US-0139899.
PR
XX 23-JUN-1999; 99US-0140353.
PR
XX 23-JUN-1999; 99US-0140354.
PR
XX 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144886.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145222.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159299.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159654.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 28.9%; Score 257; DB 21; Length 193;
Best Local Similarity 44.5%; Pred. No. 8.3e-22;
Matches 53; Conservative 20; Mismatches 32; Indels 14; Gaps 2;

OY 17 HELVPSAEADVHGLIAMLAVGSMCLFTAGFAVGGPIFPLIMSEIFPLHKGVAATGVC 76
DB 85 heavyv-----lavglmvyigsfsagmgampvwmseifpnlkvyagqma 131
OY 77 VLINMFMAFLVTKKEPNSIMELRPYGAFWLITAFCLISVLPFLITVPEPKGTLEOITA 135
DB 132 tlvnwfgawavsytlflms-wssygtfllyaalnalavfvalvpektlleqiga 189

RESULT 13
AAG37422
ID AAG37422 standard; Protein: 193 AA.
XX
AC AAG37422;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46010.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.

XX EPI033405-A2.
PN
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0125264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 28-APR-1999; 990S-0130891.
PR 30-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 04-MAY-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 06-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139452.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.

PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142820.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0152130.
PR 07-SEP-1999; 990S-0152563.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.

PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 28.9%; Score 257; DB 21; Length 193;
Best Local Similarity 44.5%; Pred. No. 8,3e-22;
Matches 53; Conservative 20; Mismatches 32; Indels 14; Gaps 2;

OY 17 HEIVPISAEPRADVHGLAMLAAGSMCLFIAGFAVGWGPFWLMLNSELPLHKGATGVC 76
Db 85 heavyv-----lavvglmvylgsfsagmgampvwmselfplnlkvgagma 131
OY 77 VLTWFMFAFLVTKFENSIMETILRPYGAFWLTAFACIIISVLTFLFVPEPTKRTIEQITA 135
Db 132 tlvwmfgawavsytlfnflms-wssygtflliyaaalalvltalvpetkxkltleqida 189

RESULT 14

AAG37398
ID AAG37398 standard; Protein; 212 AA.

XX AAG37398;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 45978.

XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX

PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 01-JUL-1999; 99US-0142154.
XX 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.

PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 28.9%; Score 257; DB 21; Length 212;
Best Local Similarity 44.5%; Pred. No. 9.3e-22;
Matches 53; Conservative 20; Mismatches 32; Indels 14; Gaps 2;

Qy 17 HELVPISEADVDHGLAVGSMCLFIAGFAVWGQIPWLMSEIFPLHIKGVATGVC 76
Db 104 heavypr-----lavvgimvyigsfsagmgampwvmseifpnlkgyvagma 150
Qy 77 VLINMFMAFLVTKEPNSIMEILRPYGAFWLTAAPCLISVLTFTLFVETGRILEQITA 135
Db 151 clvnwfgawavsytlfnflms-wssygtfllyaaInaIaIvfaIaIvpektkIeqIqa 208

RESULT 15
ACG37420 ID AAG37420 standard; Protein; 212 AA.
XX AC AAG37420;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46008.
XX
KW Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX OS
XX EPJ033405-A2.
XX PN
XX 06-SEP-2000.
XX PD
XX 25-FEB-2000; 2000EP-0301439.
PF

XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0133180.
PR 09-MAR-1999; 99US-0133548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126785.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147319.
PR 06-AUG-1999; 99US-0147560.
PR 06-AUG-1999; 99US-0147560.
PR 06-AUG-1999; 99US-0147730.
PR 09-AUG-1999; 99US-0147730.
PR 09-AUG-1999; 99US-0147735.
PR 10-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-014926.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 28.9%; Score 257; DB 21; Length 212;
Best Local Similarity 44.5%; Pred. No. 9,3e-22;

Matches 53; Conservative 20; Mismatches 32; Indels 14; Gaps 2;

OY 17 HELVPISAEPAADVHGLAMLAVGSMCLFIAGFAGVGWGPIMWLMSRIFPLHFKGVATGYC 76
Db 104 heavyv-----lavgimvylgistsagmgampwvmseIfplnlkylvagqma 150
OY 77 VLTNMFMAFLVTKFENSIIMEILRPYGAFMLTAFCLISVLFLLTFVPEFKPTLEQITA 135
Db 151 tlvmwlgawwsytlfnflms-wssygtfliyaainalaivfaivpeckkyktleqida 208

Search completed: February 13, 2002, 21:50:12
Job time: 6263 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 21:51:37 ; Search time 72.65 Seconds
(without alignments)
51.109 Million cell updates/sec

Title: US-09-516-493-10

Perfect score: 869
Sequence: 1 WRPLZNMIPRAAGIRHRLV.....VTGSPELSLMRKGVEMFQ 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/BACKFILES1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	183.5	20.6	493	2	US-09-031-392-10 Sequence 10, Appl
2	183.5	20.6	493	4	US-09-299-549-10 Sequence 10, Appl
3	179	20.1	524	2	US-08-928-692-12 Sequence 12, Appl
4	176.5	19.9	383	2	US-09-031-392-3 Sequence 3, Appl1
5	176.5	19.9	383	4	US-09-299-549-3 Sequence 3, Appl1
6	176.5	19.9	494	2	US-09-031-392-5 Sequence 5, Appl1
7	176.5	19.9	494	4	US-09-299-549-5 Sequence 5, Appl1
8	173.5	19.5	509	2	US-09-031-392-6 Sequence 6, Appl1
9	173.5	19.5	509	4	US-09-299-549-6 Sequence 6, Appl1
10	171	19.2	534	2	US-09-031-392-4 Sequence 4, Appl1
11	171	19.2	534	4	US-09-299-549-4 Sequence 4, Appl1
12	164.5	18.5	492	2	US-08-355-844-3 Sequence 3, Appl1
13	164.5	18.5	492	5	PCT-US95-16126-3 Sequence 13, Appl
14	164.5	18.5	584	2	US-08-928-692-13 Sequence 10, Appl
15	152.5	17.2	488	2	US-08-928-692-10 Sequence 11, Appl
16	122.5	13.8	488	2	US-08-928-692-11 Sequence 11, Appl
17	118.5	13.3	563	2	US-09-031-392-2 Sequence 2, Appl1
18	114.5	13.3	563	4	US-09-299-549-2 Sequence 2, Appl1
19	114.5	12.9	500	2	US-09-031-392-7 Sequence 7, Appl1
20	114.5	12.9	500	4	US-09-299-549-7 Sequence 7, Appl1
21	84.5	9.5	555	3	US-08-501-572-3 Sequence 3, Appl1
22	84.5	9.5	555	3	US-09-040-444-3 Sequence 3, Appl1
23	78.5	8.8	532	3	US-08-948-564-10 Sequence 10, Appl
24	76.5	8.6	556	3	US-08-501-572-1 Sequence 1, Appl1
25	76.5	8.6	556	3	US-09-040-444-1 Sequence 1, Appl1
26	76	8.5	322	4	US-08-964-127-6 Sequence 6, Appl1
27	74	8.3	632	1	US-08-295-814E-10 Sequence 10, Appl

28	74	8.3	632	4	US-09-343-361-10 Sequence 10, Appl
29	74	8.3	632	5	PCT-US93-01959-10 Sequence 10, Appl
30	73	8.2	520	4	US-08-964-127-2 Sequence 2, Appl1
31	72.5	8.2	817	1	US-08-381-931B-2 Sequence 2, Appl1
32	72	8.1	430	1	US-07-937-609-23 Sequence 23, Appl
33	72	8.1	430	2	US-08-919-624-3 Sequence 3, Appl1
34	72	8.1	430	4	US-08-029-170-23 Sequence 23, Appl
35	72	8.1	450	1	US-07-937-609-24 Sequence 24, Appl
36	72	8.1	450	4	US-08-029-170-24 Sequence 4, Appl1
37	72	8.1	627	1	US-08-295-814E-4 Sequence 4, Appl1
38	72	8.1	627	4	US-08-291-299-10 Sequence 10, Appl
39	72	8.1	627	4	US-09-343-361-4 Sequence 4, Appl1
40	72	8.1	627	5	PCT-US93-01959-4 Sequence 4, Appl1
41	72	8.1	627	5	PCT-US95-10579-10 Sequence 10, Appl
42	71	8.0	422	3	US-08-492-459-2 Sequence 2, Appl1
43	71	8.0	422	3	US-08-492-459-4 Sequence 4, Appl1
44	71	8.0	422	3	US-08-423-752-2 Sequence 2, Appl1
45	71	8.0	422	4	US-08-716-873-18 Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-09-031-392-10
: Sequence 10, Application US/09031392
: Patent No. 5942398
: GENERAL INFORMATION:
: APPLICANT: Tartaglia, Louis A.
: APPLICANT: Weng, Xun
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES
: TITLE OF INVENTION: ENCODING CLUTEX AND USES THEREOF
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESSES:
: ADDRESSES: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows95
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/031,392
: FILING DATE: 26-FEB-1998
: ATTORNEY/AGENT INFORMATION:
: NAME: Melkielejohn, Ph.D., Anita L.
: REGISTRATION NUMBER: 35,283
: REFERENCE/DOCKET NUMBER: 07334/072001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617/542-5070
: TELEFAX: 617/542-8906
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 493 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-031-392-10

Query Match 20.6% Score 183.5; DB 2; Length 493;
Best Local Similarity 37.7%; Pred. No. 2e-13;
Matches 40; Conservative 22; Mismatches 43; Indels 1; Gaps 1;

OY 33 LAMIAVSMCLFIAGFVAGMGPIFMILMSEIFPLHIGVATGCVLTLNWMFAPLYTKEFN 92
DB 365 MSYVIVAIFGFAVFEVGRDPIPMFTVAELFSGCPRAAIAVAGFSNMTSNFTVGLLFQ 424

APPLICANT: Hansen, Kim
TITLE OF INVENTION: Methods for Modifying the production of

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,392

ATTORNEY/AGENT INFORMATION:
NAME: Melkielejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 383 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-031-392-3

Query Match 19.9%; Score 176.5; DB 2; Length 383;
Best Local Similarity 33.3%; Pred. No. 9.3e-13;
Matches 40; Conservative 24; Mismatches 55; Indels 1; Gaps 1;

OY 33 LAMLVGSMCLFIAGFVAGCPILPMLMSIPLHLKGVATGCVLTNNFMALVTKERN 92
DB 255 MSYLSIALIFGVAFEEVGGPIPMFIVALEFSQGRPAALAVAGSNMVSNTFVGCFC 314

OY 93 SIMELIRPYGAFWLTAAFCILSVLFTLTFVPEPKGRTELOITAHLRDGDPLSVTGSPEL 152
DB 315 YVEOLCGPY-VLITFTVLLVLFITFYFKVPEPKGRTEFDEIASGFRGASQSDKTPPEL 373

RESULT 5
US-09-299-549-3
Sequence 3, Application US/09299549
Patent No. 6136547
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
TITLE OF INVENTION: GLUTEX AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299, 549
FILING DATE: 26-APR-1999
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: 09/031,392
FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Melkielejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 383 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-299-549-3

Query Match 19.9%; Score 176.5; DB 4; Length 494;
Best Local Similarity 33.3%; Pred. No. 1.3e-12;
Matches 41; Conservative 24; Mismatches 50; Indels 1; Gaps 1;

OY 19 LVPISAEPAVDHGLAMLVGSMCLFIAGFVAGCPILPMLMSIPLHLKGVATGCVLT 78
DB 348 LMTISILKKNYSMSMISICGALIVPAFELICGPIPMFIVALEFSQGRPAALAVAGSNMVSNTFVGCFC 407

OY 79 TNNFMALVTKERNSIMELIRPYGAFWLTAAFCILSVLFTLTFVPEPKGRTELOIT 134
DB 408 SNMVSNTFVGLLPPSATFYLGAV-VFIVFTVFLVIFVWTFVPEPKGRTEFDEIT 462

RESULT 7
US-09-299-549-5
Sequence 5, Application US/09299549

Query Match 19.9%; Score 176.5; DB 2; Length 494;
Best Local Similarity 33.3%; Pred. No. 1.3e-12;
Matches 41; Conservative 24; Mismatches 50; Indels 1; Gaps 1;

OY 19 LVPISAEPAVDHGLAMLVGSMCLFIAGFVAGCPILPMLMSIPLHLKGVATGCVLT 78
DB 348 LMTISILKKNYSMSMISICGALIVPAFELICGPIPMFIVALEFSQGRPAALAVAGSNMVSNTFVGCFC 407

OY 79 TNNFMALVTKERNSIMELIRPYGAFWLTAAFCILSVLFTLTFVPEPKGRTELOIT 134
DB 408 SNMVSNTFVGLLPPSATFYLGAV-VFIVFTVFLVIFVWTFVPEPKGRTEFDEIT 462

RESULT 6
US-09-031-392-5
Sequence 5, Application US/09031392
Patent No. 5942398
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES
TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,392
FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Melkielejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-031-392-5

```

? Patent No. 6136547
? GENERAL INFORMATION:
? APPLICANT: Tartaglia, Louis A.
? APPLICANT: Meng, Xun
? TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
? TITLE OF INVENTION: GLUTEX AND USES THEREOF
? NUMBER OF SEQUENCES: 10
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Fish & Richardson P.C.
? STREET: 225 Franklin Street
? CITY: Boston
? STATE: MA
? COUNTRY: USA
? ZIP: 02110-2804
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: Windows95
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/299,549
? FILING DATE: 26-APR-1999
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 09/031,392
? FILING DATE: 26-FEB-1998
? ATTORNEY/AGENT INFORMATION:
? NAME: Meiklejohn, Ph.D., Anita L.
? REGISTRATION NUMBER: 35,283
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617/542-5070
? TELEFAX: 617/542-8906
? TELEX: 200154
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 494 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-09-299-549-5

Query Match          19.9%; Score 176.5; DB 4; Length 494;
Best Local Similarity 35.3%; Pred. No. 1.3e-12;
Matches 41; Conservative 24; Mismatches 50; Indels 1; Gaps 1;

QY 19 LVPISAEADVDHGLAMLAVSMCLFTAGRAVNGPPIPLMSEIFPLHKGATGCVL 78
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 348 LMTSLILKDMYSWMSFICGAILVFAFPIGPIPIVAFELFGQPRPAMAVAGC 407
QY 79 TNNMFALVTRKFNSEIILRPYGAFWLTAFCILSVLFTLTPETKGRLEQIT 134
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 408 SNWTSNVLGLFPSPAFYIGAV-VFLVFLVFLVFWVTFKVPETKGRLEQIT 462

RESULT 8
US-09-031-392-6
? Sequence 6, Application US/09031392
? Patent No. 5942398
? GENERAL INFORMATION:
? APPLICANT: Tartaglia, Louis A.
? APPLICANT: Meng, Xun
? TITLE OF INVENTION: NUCLEIC ACID MOLECULES
? TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
? NUMBER OF SEQUENCES: 10
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Fish & Richardson P.C.
? STREET: 225 Franklin Street
? CITY: Boston
? STATE: MA
? COUNTRY: USA
? ZIP: 02110-2804
? COMPUTER READABLE FORM:
```

```

? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: Windows95
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/031,392
? FILING DATE: 26-FEB-1998
? ATTORNEY/AGENT INFORMATION:
? NAME: Meiklejohn, Ph.D., Anita L.
? REGISTRATION NUMBER: 35,283
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617/542-5070
? TELEFAX: 617/542-8906
? TELEX: 200154
? INFORMATION FOR SEQ ID NO: 6:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 509 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-09-031-392-6

Query Match          19.5%; Score 173.5; DB 2; Length 509;
Best Local Similarity 32.8%; Pred. No. 3.2e-12;
Matches 41; Conservative 26; Mismatches 55; Indels 3; Gaps 2;

QY 33 LAMLAVSGMCLFIAGFAVNGPPIPLMSEIFPLHKGATGCVL-TNNMFALVTRKFN 92
   : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 380 MSYSAIVAFIFGFAFFPIGPIPIVAFELFGQPRPAMAVAGFSNWTNFTLGGFQ 439
QY 93 SIMELIRYGAFWLTAFCILSVLFTLTPETKGRLEQITAHLRDGGDL--SYNGSP 150
   : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 440 YVAEAMGPY-VFLPAVLGLGFTLTPETKGRLEQITSAFHRTPSLDEVPKST 498
QY 151 ELSWL 155
   |||
Db 499 ELEVL 503

RESULT 9
US-09-299-549-6
? Sequence 6, Application US/09299549
? Patent No. 6136547
? GENERAL INFORMATION:
? APPLICANT: Tartaglia, Louis A.
? APPLICANT: Meng, Xun
? TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
? TITLE OF INVENTION: GLUTEX AND USES THEREOF
? NUMBER OF SEQUENCES: 10
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Fish & Richardson P.C.
? STREET: 225 Franklin Street
? CITY: Boston
? STATE: MA
? COUNTRY: USA
? ZIP: 02110-2804
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: Windows95
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/299,549
? FILING DATE: 26-APR-1999
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 09/031,392
? FILING DATE: 26-FEB-1998
? ATTORNEY/AGENT INFORMATION:
? NAME: Meiklejohn, Ph.D., Anita L.
? REGISTRATION NUMBER: 35,283
? REFERENCE/DOCKET NUMBER: 07334/072002
```

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TEXT: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 509 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-299-549-6

```

Query Match	19.5%	Score 173.5;	DB 4;	Length 509;
Best Local Similarity	32.8%	Pred. No. 3.2e-12;		
Matches 41: Conservative	26;	Mismatches 55;	Indels 3;	Gaps 2

[illegible]

RESULT 10
 US-09-031-392-4
 Sequence 4, Application US/09031392
 Patent No. 5942398
 GENERAL INFORMATION:
 APPLICANT: Tartaglia, Louis A.
 APPLICANT: Meng, Xun
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES
 TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows95
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/031,392
 FILING DATE: 26-FEB-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Melkielejohn, Ph.D., Anita L.
 REGISTRATION NUMBER: 35,283
 REFERENCE/DOCKET NUMBER: 07334/072001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 534 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-031-392-4

Query match 19.2%; Score 171; DB 2; Length 534;

Best Local Similarity 37.1%; Pred. No. 6.7e-12;
Matches 46; Conservative 17; Mismatches 53; Indels 8; Gaps 3.

```
QY      22 ISAPBADVHIGL----AMLAVGSMC---LFAGAAVGMPIMPLMSIEIPRLHKVAATG 74
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      391 ISAAMTVGLVLISOFAMMSTVSVAALFLVIFEEVSGPPIPMETIVALLSQCRRPAIA 450
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      75 VCVLTNNMFALVTAKFENSIMEILRPYGAFMLTFAFCLSVLFTLPVPETKGTLEOIT 134
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      451 VAGCGNNAACNFIVGMCFOTIADLCGP-VVVFAVALLLVFFLFAYLKVPETKKGSFEIA 509
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      135 AHLR 138
        | |
Db      510 AAFR 513
```

```

RESULT 11
US-09-299-549-4
Sequence 4, Application US/09299549
Patent No. 6136547
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
APPLICANT: Wang, Xun
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
TITLE OF INVENTION: GLUTEX AND USUS THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSD for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,549
FILING DATE: 26-Apr-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/031,392
FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Melkleyohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200134
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 534 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-299-549-4

```

Query Match	19.28;	Score 171;	DB 4;	Length 534;
Best Local Similarity	37.18;	Pred. NO. 6.7e-12;		
Matches	46;	Conservative 17;	Mismatches 53;	Indels 8;
			Gaps 3;	

Oy	22	ISAPRADHCL---	AMLAVGSMC-LF	IAGAVGSMGP	PIYMLSEIRPLN	KIVATG	74		
Db	391	ISAAVMTGLTSSQ	FAMSVSVVALE	IFVIFEEVGPR	IPMTVALE	LSQGRPAIA	450		
Oy	75	VCULTNFMAL	LYTKERN	IMELLRP	IGAWLTA	FCILSVLPT	IVPRTKRTLEQIT	134	
Db	451	VAGFCNMACN	IVDMCROY	IADLCGR	VFVAVFV	VLILFF	FLAYLKV	VPETKKSFEIA	509

Thu Feb 14 07:43:57 2002

us-09-516-493-10.ra1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 21:53:10 : Search time 79.58 Seconds

(without alignments)
157,939 Million cell updates/sec

Title: US-09-516-493-10

Perfect score: 889

Sequence: 1 WRPLZNMWIPRAGIRHELY.....VTGSPELSLWRVSKGVEMPQ 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.*
2: pir1.*
3: pir3.*
4: pir4.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	257	28.9	348	2	D84922
2	225	25.3	487	2	E96782
3	213.5	24.0	490	2	T14545
4	212	23.8	457	2	E70070
5	209.5	23.6	483	2	D96589
6	204.5	23.0	516	2	T12199
7	202	22.7	521	2	G84864
8	198.5	22.3	522	2	S12042
9	198.5	22.3	522	2	E86246
10	197.5	22.2	491	2	A26430
11	197.5	22.2	491	2	B86096
12	195	21.9	523	2	S25015
13	194.5	21.9	378	2	G86220
14	194.5	21.9	523	2	T10122
15	194	21.8	493	2	F66696
16	193.5	21.8	461	2	G85059
17	193.5	21.8	582	2	F71431
18	193	21.7	474	2	B86221
19	193	21.7	496	2	T52132
20	193	21.7	734	2	H63440
21	192.5	21.7	534	2	S88435
22	192.5	21.7	580	2	D86426
23	191	21.6	523	2	S06920
24	191	21.5	472	2	S47089
25	190.5	21.4	534	2	S14144
26	190	21.4	467	2	B96829
27	190	21.4	472	2	B26430
28	190	21.4	472	2	E85936
29	188.5	21.2	496	2	A31986

30	188	21.1	490	2	F86220	hypothetical prote
31	188	21.1	665	2	T51198	related to myo-ino
32	185.5	20.9	514	2	S25009	monosaccharide tra
33	185.5	20.9	549	2	T14606	probable sugar tra
34	185	20.8	522	2	A31556	glucose transport
35	184.5	20.8	209	2	T14798	hypothetical prote
36	184.5	20.8	464	2	F65079	galactose-proton s
37	184.5	20.8	729	2	T06127	probable sugar tra
38	183	20.6	461	2	D70073	metabolite transpo
39	182	20.5	508	2	T05156	probable glucose t
40	182	20.5	606	2	T27072	hypothetical prote
41	181.5	20.4	493	2	A41751	glucose transport
42	181	20.4	547	2	C84593	probable sugar tra
43	180.5	20.3	464	2	F85951	galactose-proton s
44	180	20.2	580	2	D84772	probable sugar tra
45	179	20.1	524	2	A31318	glucose transporte

ALIGNMENTS

```
RESULT 1
D84922
Probable sugar transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: D84922
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.: Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, A.M.; VanAken, S.E.; Umeyam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; M0ID:20083487
A:Accession: D84922
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-348 <STO>
A:Cross-References: GB:A8002093; NID:g4249409; PIDN:AAD13706.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g48020
A:Map position: 2

Query Match 28.9%; Score 257; DB 2; Length 348;
Best Local Similarity 44.5%; Pred. No. 9.8e-17;
Matches 53; Conservative 20; Mismatches 32; Indels 14; Gaps 2;

QY 17 HELVPSAEPPADYHLGLAVGSMCLFIAGFAVGWGPITWLMSEIFPHIKGVATGVC 76
      ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| :
Db 240 HEAVPV-----LAVVGIMVYICFSAGMGAMPVWVSEIFPINIKVAGGMA 286

QY 77 VLTNMFAPLVTKEFNSIMELIRPYGAFWLTAFCILSVLFTLFVEETGRTLEQTA 135
      | | | | : : : : : | | | | : : : : : | | | | : : : : : |
Db 287 TLVNWEGAMAVSTFENFLMS-WSSYGFLLIYAAINLAIVFAIVPEYTKGTLEQIDA 344

RESULT 2
E96782
hypothetical protein F22H5.6 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: E96782
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
ansen, N.P.; Hughes, B.; Huizart, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talio
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; M0ID:21016719
```

h, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hul

```

03 EVINAEFNSIMEIKRIGAFWLTIAAFCLISVLEFLTVFVPETRGRTLEQI 133
   :::: | | :: | | :: | | :: | | :: | | :: | | :: | | :: | |
427 WTCGEMENRIMR EVMSGRTUWLLTLLTLLTLLTLLTLLTLLTLLTLLTLLT

```

Db 427 WISTFENLMN-WNPAGTFYVEATVCGATVIFVAKLVPEFGRTLEEL 474

RESULT 6
T12199
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 24-Sep-1999
C:Accession: S12042; S14627
R:Sauer, N.; Friedlaender, K.; Graeme-Wicke, U.
EMBO J. 9, 3045-3050, 1990
A:Title: Primary structure, genomic organization and heterologous expression of a glu
A:Reference number: S12042; MUID:91005995
A:Accession: S12042
A:Molecule type: mRNA
A:Residues: 1-522 <STO>
A:Cross-references: EMBL:555350; NID:g16519; PIDN:CAA3907.1; PID:g16520
A:Note: The sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 459-
C:Genetics:
A:Gene: SRP1
C:Superfamily: glucose transport protein
C:Keywords: sugar transport; transmembrane protein

Query Match 23.0%; Score 204.5; DB 2; Length 516;
Best Local Similarity 37.8%; Pred. No. 1.2e-11;
Matches 42; Conservative 24; Mismatches 44; Indels 1; Gaps 1;
OY 23 SAEPADVHLGLAMLVGSMCLFIAGFAGVGMGPFWLMLSEIFPLHKGATGVCVLTNNF 82
DB 376 SCEPDLPKRWAIIVVTFICIVAGFAMSGPLGMLVPSSEIFPLEIRSAQSITVSVM 435
OY 83 MAFLVTKFNSIMELRPYGAFMLTAACILSVLFTLPVETKGRLEOI 133
DB 436 FTFLVAQIFLTMCHMK-FGLFLVFAFFVVMISFYIFLPETKGIPIDEM 485

RESULT 7
G84864
probable membrane transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: G84864
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Molitz, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: G84864
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-521 <STO>
A:Cross-references: GB:AE002093; NID:g2289003; PIDN:AAB64332.1; GSPDB:GN00139
C:Genetics:
A:Gene: AT2g43330
A:Map position: 2
C:Superfamily: glucose transport protein

Query Match 22.7%; Score 202; DB 2; Length 521;
Best Local Similarity 37.4%; Pred. No. 2.2e-11;
Matches 37; Conservative 19; Mismatches 43; Indels 0; Gaps 0;
OY 35 WLAVGSMCLFIAGFAGVGMGPFWLMLSEIFPLHKGATGVCVLTNNFMAFLVTKFENS 94
DB 389 WLAVGLALIVFAPCGPVPMTVNSEIYPQYRGICGMSATVNMISNLIVAOFTLTI 448
OY 95 MEILRPYGAFMLTAACILSVLFTLPVETKGRLEOI 133
DB 449 AEAAGTGMFTFLAGIAVLAVIVFVPEOTGLTFSEV 487

RESULT 8
S12042
glucose transport protein SRP1 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 24-Sep-1999
C:Accession: S12042; S14627
R:Sauer, N.; Friedlaender, K.; Graeme-Wicke, U.
EMBO J. 9, 3045-3050, 1990
A:Title: Primary structure, genomic organization and heterologous expression of a glu
A:Reference number: S12042; MUID:91005995
A:Accession: S12042
A:Molecule type: mRNA
A:Residues: 1-522 <STO>
A:Cross-references: EMBL:555350; NID:g16519; PIDN:CAA3907.1; PID:g16520
A:Note: The sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 459-
C:Genetics:
A:Gene: SRP1
C:Superfamily: glucose transport protein
C:Keywords: sugar transport; transmembrane protein

Query Match 22.3%; Score 198.5; DB 2; Length 522;
Best Local Similarity 33.9%; Pred. No. 4.6e-11;
Matches 38; Conservative 28; Mismatches 45; Indels 1; Gaps 1;
OY 22 ISAEPPADVHLGLAMLVGSMCLFIAGFAGVGMGPFWLMLSEIFPLHKGATGVCVLTNNF 81
DB 375 VDSPEELPKRWAIIVVTFICIVAGFAMSGPLGMLVPSSEIFPLEIRSAQSITVSVM 434
OY 82 MAFLVTKFNSIMELRPYGAFMLTAACILSVLFTLPVETKGRLEOI 133
DB 435 IFFTLIAQIFLTMCHMK-FGLFLVFAFFVVMISFYIFLPETKGIPIDEM 485

RESULT 9
E86246
glucose transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: E86246
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Xu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: E86246
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-522 <STO>
A:Cross-references: GB:AE005172; NID:g5734730; PIDN:AAD49995.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
C:Superfamily: glucose transport protein

Query Match 22.3%; Score 198.5; DB 2; Length 522;
Best Local Similarity 33.9%; Pred. No. 4.6e-11;
Matches 38; Conservative 28; Mismatches 45; Indels 1; Gaps 1;
OY 22 ISAEPPADVHLGLAMLVGSMCLFIAGFAGVGMGPFWLMLSEIFPLHKGATGVCVLTNNF 81
DB 375 VDSPEELPKRWAIIVVTFICIVAGFAMSGPLGMLVPSSEIFPLEIRSAQSITVSVM 434
OY 82 MAFLVTKFNSIMELRPYGAFMLTAACILSVLFTLPVETKGRLEOI 133
DB 435 IFFTLIAQIFLTMCHMK-FGLFLVFAFFVVMISFYIFLPETKGIPIDEM 485

RESULT 10
A26430

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: G86220
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-378 <STO>
 A:Cross-references: GB:AE005172; NID:g2342668; PIDN:AAB70414.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 21.9%; Score 194.5; DB 2; Length 378;
 Best Local Similarity 35.1%; Pred. No. 7.9e-11;
 Matches 33; Conservative 29; Mismatches 31; Indels 1; Gaps 1;

QY 44 FICGFAVGMGPIMLMSSEIPPLHIKGVATGCVLTMMFMFLYTKRNSIMELIRPYGA 103
 DB 282 YVSSFGIGLGLPWIIMSEIFPVNWKITAGSLVYMSNMFMIITYSFEMIQ-WSASGT 340

QY 104 FMULTAFCLSVLFTLTFVPEPTKGRTEQITAH 137
 DB 341 YFIFSGVSLVTIVFIMTLVPEPTKGRTELEIOTSL 374

RESULT 14
 T10122

hexose transport protein STC - castor bean

C:Species: Ricinus communis (castor bean)

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T10122

R:Weig, A.; Franz, J.; Sauer, N.; Komor, E.

A:Title: Isolation of a family of cDNA-clones from Ricinus communis L. with close homolo

A:Reference number: 216961

A:Accession: T10122

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-523 <MEI>

A:Cross-references: EMBL:L08196; NID:g169717; PID:g169718

A:Experimental source: strain Carmencita; tissue-type cotyledon

C:Genetics:

A:Gene: STC

C:Superfamily: glucose transport protein

C:Keywords: sugar transport; transmembrane protein

Query Match 21.9%; Score 194.5; DB 2; Length 523;
 Best Local Similarity 36.6%; Pred. No. 1.1e-10;
 Matches 41; Conservative 22; Mismatches 48; Indels 1; Gaps 1;

QY 22 ISAEPAVHGLAMLAAGSMCLFIAGFAGVGMGPIMLMSSEIPPLHIKGVATGCVLTNM 81
 DB 377 VDCAPGDLPMYAVVYVLFICIVSGFAMSGPLGMLPSEIPLETIRSAOSVNSVNM 436

QY 82 FMAFLVTKRNSIMELIRPYGAPWLTAFCLISVLTFTLVPEPTKGRTEQIT 133
 DB 437 FTFEVAQVF-LIMLCHLKGLFIFSFVILMSIFVYFLPRTKGIPIEM 487

RESULT 15
 F96696

Protein PIN21.12 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C:Accession: F96696

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitli, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: F96696

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-493 <STO>

A:Cross-references: GB:AE005173; NID:g9828628; PIDN:AAG00251.1; GSPDB:GN00141

C:Genetics:

A:Gene: PIN21.12

A:Map position: 1

C:Superfamily: glucose transport protein

Query Match 21.8%; Score 194; DB 2; Length 493;
 Best Local Similarity 37.0%; Pred. No. 1.2e-10;
 Matches 40; Conservative 22; Mismatches 46; Indels 0; Gaps 0;

QY 30 HUGLAMLAVGSMCLFIAGFAGVGMGPIMLMSSEIPPLHIKGVATGCVLTMMFMFLYTK 89
 DB 382 HPSALCLSVGTLVFLTFALGAGVPGLLPEIPSRIRAKAMAFCSVHWVINFVGL 441

QY 90 FENSMELIRPYGAPWLTAFCLISVLTFTLVPEPTKGRTEQITAH 137
 DB 442 LFLRLERLGPRLYSMEFTCLMAVMEVKRVETKGRTEQITAH 489

Search completed: February 13, 2002, 21:53:11
 Job time: 6133 sec

Thu Feb 14 07:44:03 2002

us-09-516-493-10.rpr

Match	Local Similarity	Best Local Similarity	Conservative	Mismatch	Score	DB	Length
44	37.38	37.38	44	32	197.5	1	491

FT TRANSMEM 309 329 POTENTIAL.
 FT TRANSMEM 342 362 POTENTIAL.
 FT CONFLICT 400 401 RP -> SA (IN REF. 3).
 SQ SEQUENCE 461 AA; 50235 MW; B6587B5DC7272EF1 CRC64;

Query Match 21.8%; Score 194; DB 1; Length 461;
 Best Local Similarity 37.5%; Pred. No. 7.6e-10;
 Matches 45; Conservative 17; Mismatches 38; Indels 20; Gaps 4;

OY 31 LGT-----AMLAAGSMCLFAGFAGVAGCPIMWLMSEFLPHITGATGCVLTNNMFMAFL 86
 ||| | : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 331 LGTSASATAMTGVFLGVITVATWGPVWMLPELFPKRAATG-----FTTL 382
 OY 87 VTKFENSIMEI-----LRPYGAFMLTAAF-----CILSVLFLTFVPEPTKGRTEQITAHIR 138
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :
 DB 383 VLSAAMLIVSLVPLMLRPMGIAMVPMVSVICLLSFFFAFYVNPETKKSLEIEASLK 442

RESULT 5
 HUP3_CHLKE STANDARD: PRT; 534 AA.

ID HUP3_CHLKE STANDARD: PRT; 534 AA.
 AC Q39525;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE H(+)/HEXOSE COTRANSPORTER 3.
 GN HUP3.
 OS Chlorella kessleri.
 OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorella.
 OX NCBI_TaxID=3074;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95175627; PubMed=7870840;
 RA Steidler R, Wolf K, Hilgath C, Tanner W, Sauer N.K.;
 RT "Subcellular localization of the inducible Chlorella Hup1
 symporter.";
 RT moscarcharide-H+ symporter and cloning of a Co-induced galactose-H+
 symporter.";
 RL Plant Physiol. 107:33-41(1995).
 CC -1- FUNCTION: ACTIVE UPTAKE OF HEXOSES.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X75440; CAA53192.1; -
 DR InterPro: IPR003663; Sugar trnsportr.
 DR Pfam: PF00083; sugar.tr.1
 DR PRINTS: PR00171; SUGRTRANSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane; Transport; Sugar transport; Symport; Repeat;
 KW Multigene family.
 FT TRANSMEM 21 41 POTENTIAL.
 FT TRANSMEM 87 107 POTENTIAL.
 FT TRANSMEM 114 134 POTENTIAL.
 FT TRANSMEM 137 157 POTENTIAL.
 FT TRANSMEM 172 192 POTENTIAL.
 FT TRANSMEM 205 225 POTENTIAL.
 FT TRANSMEM 306 326 POTENTIAL.
 FT TRANSMEM 329 349 POTENTIAL.
 FT TRANSMEM 353 373 POTENTIAL.
 FT TRANSMEM 395 415 POTENTIAL.
 FT TRANSMEM 431 451 POTENTIAL.
 FT TRANSMEM 456 476 POTENTIAL.
 SQ SEQUENCE 534 AA; 57772 MW; 52E03715F3726017 CRC64;

Query Match 21.7%; Score 192.5; DB 1; Length 534;
 Best Local Similarity 38.2%; Pred. No. 1.2e-09;
 Matches 42; Conservative 25; Mismatches 38; Indels 5; Gaps 3;

OY 26 PADVHLGLAMLAAGSMCLFAGFAGVAGCPIMWLMSEFLPHITGATGCVLTNNMFMAFL 85
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :
 DB 386 PHFVSAGV--LAV--LCITAGFAMSGPMGLIPSEITLLETRPAGTAVAAMGNLFSE 441
 OY 86 LVTKFENSIMEI-----LRPYGAFMLTAAF-----CILSVLFLTFVPEPTKGRTEQITAHIR 135
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :
 DB 442 VLGQAFVSMCAAK--FGVFLFFAGMLVIMVLCALFLDETETKGVPIERVOA 490

RESULT 6
 ARAE_KLEOX STANDARD: PRT; 472 AA.

ID ARAE_KLEOX STANDARD: PRT; 472 AA.
 AC P45598;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ARABINOSE-PROTON SYMPORTER (ARABINOSE TRANSPORTER).
 GN ARAE.
 OS Klebsiella oxytoca.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Klebsiella.
 OX NCBI_TaxID=571;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8017;
 RX MEDLINE=95394866; PubMed=7665532;
 RA Shatwell K.P., Charalambous B.M., McDonald T.P., Henderson P.J.F.;
 RT "Cloning, sequencing, and expression of the araE gene of Klebsiella
 oxytoca 8017, which encodes arabinose-H+ symport activity.";
 RL J. Bacteriol. 177:5379-5380(1995).
 CC -1- FUNCTION: UPTAKE OF ARABINOSE ACROSS THE BOUNDARY MEMBRANE WITH
 CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X79598; CAA56110.1; -
 DR InterPro: IPR003663; Sugar trnsportr.
 DR Pfam: PF00083; sugar.tr.1
 DR PRINTS: PR00171; SUGRTRANSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport.
 FT TRANSMEM 23 43 POTENTIAL.
 FT TRANSMEM 63 83 POTENTIAL.
 FT TRANSMEM 92 110 POTENTIAL.
 FT TRANSMEM 120 138 POTENTIAL.
 FT TRANSMEM 149 167 POTENTIAL.
 FT TRANSMEM 177 197 POTENTIAL.
 FT TRANSMEM 259 278 POTENTIAL.
 FT TRANSMEM 299 318 POTENTIAL.
 FT TRANSMEM 328 348 POTENTIAL.
 FT TRANSMEM 362 382 POTENTIAL.
 FT TRANSMEM 396 415 POTENTIAL.
 FT TRANSMEM 424 443 POTENTIAL.
 SQ SEQUENCE 472 AA; 51732 MW; 410021E1BEE3D96E CRC64;

Query Match 21.5%; Score 191; DB 1; Length 472;

RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Toke T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shingawa H.,
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 RN [6]
 RP PRELIMINARY SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE=8411468; PubMed=6319708;
 RA Storer C., Schleif R.F.,
 RT "The araE low affinity L-arabinose transport promoter. Cloning,
 RT sequence, transcription start site and DNA binding sites of
 RT regulatory proteins.";
 RL J. Mol. Biol. 171:369-381(1983).
 CC -1- FUNCTION: UPTAKE OF ARABINOSE ACROSS THE BOUNDARY MEMBRANE WITH
 CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: J03732; AAA23469.1; -
 DR EMBL: X00272; CAA25075.1; -
 DR EMBL: U29581; ABA04048.1; -
 DR EMBL: AE000368; AAC75880.1; -
 DR EMBL: AE005513; AAG57953.1; -
 DR EMBL: AP002563; BAB37121.1; -
 DR PIR: B26430; B26430.
 DR PIR: A28075; A28075.
 DR Ecocore: EG10056; araE.
 DR InterPro: IPR003663; Sugar_trnsport.
 DR InterPro: IPR003662; sub_trnsport.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGARTRANSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
 KW Complete proteome.
 FT TRANSMEM 23 43 POTENTIAL.
 FT TRANSMEM 63 83 POTENTIAL.
 FT TRANSMEM 92 110 POTENTIAL.
 FT TRANSMEM 120 138 POTENTIAL.
 FT TRANSMEM 149 167 POTENTIAL.
 FT TRANSMEM 177 197 POTENTIAL.
 FT TRANSMEM 259 278 POTENTIAL.
 FT TRANSMEM 299 318 POTENTIAL.
 FT TRANSMEM 328 348 POTENTIAL.
 FT TRANSMEM 362 382 POTENTIAL.
 FT TRANSMEM 396 415 POTENTIAL.
 FT TRANSMEM 424 443 POTENTIAL.
 FT CONFLICT 26 28 SVA -> YDR (IN REF. 6).
 SO SEQUENCE 472 AA: 51684 MW: 41190044ID44393 CRC64;

Query Match 21.4%; Score 190; DB 1; Length 472;
 Best Local Similarity 36.4%; Pred. No. 1.7e-09;
 Matches 40; Conservative 17; Mismatches 53; Indels 0; Gaps 0;

OY 32 GLAMTAVSGMCLFAGFAVAGCPDIPWLLMSEIFPLHIKGVATGCVLTWFMFAFLVTKPEF 91
 DB 356 GLSWLSVGMTCMTCIAGYASAPVWVILCEIOPLKCRDGFJCTCTTWMVSNMIGATF 415
 OY 92 NSIMELRLRYGAFWLTACFIIISVLTFLTFVETPKRTLEQTAAHRRDD 141
 DB 416 LTLDSIGAAGTFWLTALNIAFVGITFWLIPETKNVLEHIERKLAGE 465

RESULT 9
 CTR2_MOUSE
 ID GTR2_MOUSE STANDARD: PRT: 523 AA.
 AC P14246;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE GLUCOSE TRANSPORTER TYPE 2, LIVER.
 GN SLC2A2 OR GLUT2 OR GLUT-2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALE/C; TISSUE=Liver;
 RX MEDLINE=90098776; PubMed=2602116;
 RA Suzue K., Lodish H.F., Thorens B.,
 RT "Sequence of the mouse liver glucose transporter.";
 RT Nucleic Acids Res. 17:10095-10099(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Liver;
 RX MEDLINE=89366666; PubMed=2771649;
 RA Asano T., Shimasaki Y., Lin J.L., Akanuma Y., Takaku F., Oka Y.,
 RT "The nucleotide sequence of cDNA for a mouse liver-type glucose
 RT transporter protein.";
 RT Nucleic Acids Res. 17:6386-6386(1989).
 RN [3]
 RP SEQUENCE OF 384-496 FROM N.A.
 RX MEDLINE=92111400; PubMed=1765007;
 RA Hogen A., Heyner S., Charbon M.J., Copeland N.G., Gilbert D.J.,
 RA Jenkins N.A., Thorens B., Schultz G.A.,
 RT "Glucose transporter gene expression in early mouse embryos.";
 RT Development 113:363-372(1991).
 RN [4]
 RP TISSUE SPECIFICITY.
 RC STRAIN=C57BL/6;
 RX MEDLINE=93170163; PubMed=1289053;
 RA Smith D.E., Gridley T.,
 RT "Differential screening of a PCR-generated mouse embryo cDNA library:
 RT glucose transporters are differentially expressed in early
 RT postimplantation mouse embryos.";
 RT Development 116:555-561(1992).
 CC -1- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY
 CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
 CC MEMBRANE OF HEPATOCYTES & IS RESPONSIBLE FOR UPTAKE OF GLUCOSE BY
 CC THE BETA CELLS; IT MAY COMPRISE PART OF THE GLUCOSE-SENSING
 CC MECHANISM OF THE BETA CELL. IT MAY ALSO PARTICIPATE WITH THE
 CC NA(+) /GLUCOSE COTRANSPORTER IN THE TRANSCYLLULAR TRANSPORT OF
 CC GLUCOSE IN THE SMALL INTESTINE & KIDNEY.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: IN EMBRYO, EXPRESSED IN ENDODERM LAYER OF
 CC YOLK SAC AND LIVER PRIMORDIUM.
 CC -1- SIMILARITY: HIGH TO OTHER GLUCOSE TRANSPORTERS.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X16986; CAA34855.1; -
 DR EMBL: X15684; CAA33719.1; -
 DR EMBL: S7926; AAB20847.1; -
 DR PIR: S06920; S06920.
 DR PIR: S05319; S05319.
 DR MGI: 1095438; Slc2a2.

[1]
SEQUENCE FROM N.A.
RX MEDLINE=99003066; PubMed=3048704;
RA Thorens B., Sarkar H.K., Kaback H.R., Lodish H.F.;
RT "Cloning and functional expression in bacteria of a novel glucose
transporter present in liver, intestine, kidney, and beta-pancreatic
islet cells."
RL Cell 55:281-290(1988).
CC -1- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY
CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
CC MEMBRANE OF HEPATOCYTES & IS RESPONSIBLE FOR UPTAKE OF GLUCOSE BY
CC THE BETA CELLS. IT MAY COMPRISE PART OF THE GLUCOSE-SENSING
CC MECHANISM OF THE BETA CELL. IT MAY ALSO PARTICIPATE WITH THE
CC NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCELLULAR TRANSPORT OF
CC GLUCOSE IN THE SMALL INTESTINE & KIDNEY.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: PRESENT IN LIVER, INTESTINE, KIDNEY AND
CC BETA-PANCREATIC ISLET CELLS.
CC -1- SIMILARITY: HIGH TO OTHER GLUCOSE TRANSPORTERS.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: J03145; AAA41298.1; -
DR PIR: A31556; A31556.
DR InterPro: IPR003663; Sugar_tnsport.
DR InterPro: IPR003662; sub_tnsport.
DR Pfam: PF00083; sugar_tr_1.
DR PRINTS: PR00171; SUGRTNSPORT.
DR PRINTS: PR00172; GLUCTRNSPORT.
DR PRINTS: PR01191; GLUCTRNSPORT2.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Transport; Glycoprotein;
KW Multigene family.
FT TRANSSEM 11 31
FT DOMAIN 32 97
FT TRANSSEM 98 117
FT TRANSSEM 126 146
FT TRANSSEM 157 177
FT TRANSSEM 186 206
FT TRANSSEM 216 236
FT DOMAIN 237 301
FT TRANSSEM 302 322
FT TRANSSEM 337 357
FT TRANSSEM 368 388
FT TRANSSEM 399 419
FT TRANSSEM 432 452
FT TRANSSEM 460 480
FT DOMAIN 481 522
KW CARBOHYD
SEQUENCE 522 AA: 57085 MW: 075AB81E56CF33F7 CRC64;

Query Match 20.8%; Score 185; DB 1; Length 522;
Best Local Similarity 34.7%; Pred. No. 5.2e-09;
Matches 42; Conservative 21; Mismatches 48; Indels 10; Gaps 3;

OY 31 LGL-----AWLAVGSMC---LFIAGFVAGMGPIMLMSSEIFPLHIGVATGCVLTNW 81
DB 383 LGLVLDKFTWMSYVSMALFLFVSFFELGPGPLPWEVADEFSGCPPTALAAAFSNW 442
OY 82 FMALVTEPNSIMELRPYGFMTAAFCILSVLFTLPEVETGRLTQALHRLRGD 141
DB 443 VCNFTIALCFQYIADFLGPY-VFELFAGVAVLFTLTFKXVPEKTKGKSFDDIAAEFRKS 501
OY 142 G 142

DB 502 G 502
RESULT 12
ID GALP_ECOLI STANDARD: PRT; 464 AA.
AC P37021;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE GALACTOSE-PROTON SYMPORTER (GALACTOSE TRANSPORTER).
GN GALP OR B2943.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
[1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RA Roberts P.E.;
RL Thesis (1992), University of Cambridge, U.K.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatter F.R., Plunkett G., II, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -1- FUNCTION: UPTAKE OF GALACTOSE ACROSS THE BOUNDARY MEMBRANE WITH
CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: U28377; AAA69110.1; -
DR EMBL: AE000377; AAC75980.1; -
DR Ecocore: E612148; galp.
DR InterPro: IPR003663; Sugar_tnsport.
DR InterPro: IPR003662; sub_tnsport.
DR Pfam: PF00083; sugar_tr_1.
DR PRINTS: PR00171; SUGRTNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
KW Complete proteome.
FT TRANSSEM 16 36
FT TRANSSEM 57 77
FT TRANSSEM 85 105
FT TRANSSEM 113 133
FT TRANSSEM 140 160
FT TRANSSEM 172 192
FT TRANSSEM 251 271
FT TRANSSEM 291 311
FT TRANSSEM 322 342
FT TRANSSEM 352 372
FT TRANSSEM 395 415
FT TRANSSEM 417 437
KW
SEQUENCE 464 AA: 50982 MW: 07E08935B8E3F8E CRC64;

Query Match 20.8%; Score 184.5; DB 1; Length 464;
Best Local Similarity 35.0%; Pred. No. 5.1e-09;
Matches 41; Conservative 20; Mismatches 51; Indels 5; Gaps 1;

OY	29	VHLG3-----AMLAGSCKPIAGAVCGCPIMPLMSIEPFLHJKVATGCVCITNFM	83
	:	: : :	:
Dd	338	MHICGHSSAOYFALAMLMFLVGFAMASGPIILWLCSEIOPCLKKRDGCTTSTATNMIA	397
	:	: : :	:
OY	84	AFVLTKRPNISIMEILRPYCAPMLTAFCILSLVLTLPVEPKGRLEDPITAHLRDG	140
	:	: : :	:
Dd	398	NMIYGATEFTLMTLNTGNANTFNVVALMVLFIILTLLMLVPETKHVSLEHIBENLMKG	454
	:	: : :	:
RESULT	13		
GTR3_CANFA		STANDARD:	PRT: 495 AA.
ID	GTR3_CANFA		
AC	P47842:		
DT	01-FEB-1996 (Rel. 33, Created)		
DT	01-FEB-1996 (Rel. 33, Last sequence update)		
DT	15-JUL-1999 (Rel. 38, Last annotation update)		
DE	GLUCOSE TRANSPORTER TYPE 3, BRAIN.		
GN	SLOC2A3 OR GLUT3.		
OS	Canis familiaris (Dog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OX	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
RN	NCBI_TaxID=9615;		
RP	(1)		
RC	SEQUENCE FROM N.A.		
RX	TISSUE=Brain cortex;		
RA	MEDLINE=66194907; PubMed=8654954;		
RT	Borsion N.D., Salio W.L., Drewes L.R.;		
RL	"Canine brain glucose transporter 3: gene sequence, phylogenetic comparisons and analysis of functional sites.";		
CC	Gene 168:251-256(1996).		
CC	-1- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. PROBABLY A NEURONAL		
CC	GLUCOSE TRANSPORTER.		
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.		
CC	-1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN.		
CC	-1- SIMILARITY: HIGH TO OTHER GLUCOSE TRANSPORTERS.		
CC	-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isdb.ch/announce/ or send an email to license@isdb-sib.ch).		
CC	-----		
DR	EMBL, L35267; AAA51454.1; -		
DR	InterPro: IPR003663; Sugar_tlnsportr.		
DR	InterPro: IPR003662; sub_tlnsportr.		
DR	Pfam: PF00083; sugar_tr.1.		
DR	PRINTS: PR00171; SUGSTRNSPORT.		
DR	PRINTS: PR00172; GLUCTRNSPORT.		
DR	PRINTS: PR01192; GLUCTRNSPORT3.		
DR	PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.		
DR	PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.		
KW	Transmembrane; Sugar transport; Transport; Glycoprotein;		
KW	Multigene family.		
FT	TRANSMEM	11	31 POTENTIAL.
FT	DOMAIN	32	64 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	65	85 POTENTIAL.
FT	TRANSMEM	94	114 POTENTIAL.
FT	TRANSMEM	125	145 POTENTIAL.
FT	TRANSMEM	154	174 POTENTIAL.
FT	TRANSMEM	184	204 POTENTIAL.
FT	DOMAIN	205	269 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	270	290 POTENTIAL.
FT	TRANSMEM	305	325 POTENTIAL.
FT	TRANSMEM	336	355 POTENTIAL.
FT	TRANSMEM	367	387 POTENTIAL.
FT	TRANSMEM	400	420 POTENTIAL.
FT	TRANSMEM	428	448 POTENTIAL.
FT	DOMAIN	449	495 CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	43	43 N-LINKED (GLCNAC...) (POTENTIAL).

```

SQ SEQUENCE 495 AA; 54283 MW; 09063C013DAE39C3 CRC64;

Query Match      20.5%; Score 182; DB 1; Length 495;
Best Local Similarity 34.4%; Pred. No. 9e-09;
Matches 45; Conservative 26; Mismatches 44; Indels 16; Gaps 3;

OY 19 LVPISAEADPAVHLGLAMVLAVGSMCLFIQFANGKGPPIWLMSELPPLHIKVAATGYCVL 78
   || | : : : : : || | : : : : : || | : : : : : || | : : : : : ||
DB 348 IMTIISLLKLDNWNMSFCVIGALIVFAFEETGPDPIMFIVAELFSQGRPAAAMVAGC 407
OY 79 TWMPAFAPLVTKRDFNSIMETLRPYCAFMLTA-----AFCILSVLTFLTFPETKGTLE 131
   || | | | | : : | | | | : : | | | | : : | | | | : : | | | | : : |
DB 408 SMMTSGNFVL-----GLTPPSAAFYLCAAYFIITFGFLIVFLVFPPKVPETKGRFFE 459
OY 132 QITAHLRDDCG 142
   || : : : : : ||
DB 460 ELTRAF-EGCG 469

RESULT# 14
GTR3_MOUSE
ID GTR3_MOUSE STANDARD: PRT; 493 AA.
AC P32037.
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE GLUCOSE TRANSPORTER TYPE 3, BRAIN.
GN SLC2A3 OR GLUT3 OR GLUT-3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mamalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=92112695; PubMed=1730609;
RA Nadamatsu S., Kornhauser J.M., Seino S., Mayo K.E., Steiner D.F.,
RA Bell G.I.;
RT "Glucose transporter expression in brain. cDNA sequence of mouse
RT GLUT3, the brain facilitative glucose transporter isoform, and
RT identification of sites of expression by in situ hybridization.";
RL J. Biol. Chem. 267:467-472(1992).
[2]
RN SEQUENCE FROM N.A.
RP Takeda J., Minokoshi Y., Yasuda K., Kayano T., Graeme B.I.;
RL Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. PROBABLY A NEURONAL
CC GLUCOSE TRANSPORTER.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN.
CC -!- SIMILARITY: HIGH TO OTHER GLUCOSE TRANSPORTERS.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
-----
DR EMBL; M75135; AAA37704.1; -.
DR EMBL; X61093; CAA43406.1; -.
DR EMBL; U11853; AAB60666.1; -.
DR EMBL; U11844; AAB60666.1; JOINED.
DR EMBL; U11845; AAB60666.1; JOINED.
DR EMBL; U11846; AAB60666.1; JOINED.
DR EMBL; U11848; AAB60666.1; JOINED.
DR EMBL; U11849; AAB60666.1; JOINED.
DR EMBL; U11850; AAB60666.1; JOINED.
DR EMBL; U11851; AAB60666.1; JOINED.
DR EMBL; U11852; AAB60666.1; JOINED.
```

DR PIR: A41751; A41751.
 DR MGD: MGI:95757; Slic2a3.
 DR Interpro: IPR003663; Sugar_transport.
 DR Interpro: IPR003662; sub_transport.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGTRANSPORT.
 DR PRINTS: PR00172; GLUCTRSPORT.
 DR PRINTS: PR01192; GLUCTRSPORT3.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 DR Transmembrane: Sugar transport; Transport; Glycoprotein;
 KM Multigene family.
 FT TRANSMEM 11 31 POTENTIAL.
 FT DOMAIN 32 64 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 65 85 POTENTIAL.
 FT TRANSMEM 94 114 POTENTIAL.
 FT TRANSMEM 125 145 POTENTIAL.
 FT TRANSMEM 154 174 POTENTIAL.
 FT TRANSMEM 184 204 POTENTIAL.
 FT DOMAIN 205 269 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 270 290 POTENTIAL.
 FT TRANSMEM 305 325 POTENTIAL.
 FT TRANSMEM 336 355 POTENTIAL.
 FT TRANSMEM 367 387 POTENTIAL.
 FT TRANSMEM 400 420 POTENTIAL.
 FT TRANSMEM 428 448 POTENTIAL.
 FT DOMAIN 449 493 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 493 AA; 53478 MW; 90908BDCB8780082 CRC64;

Query Match 20.4%; Score 181.5; DB 1; Length 493;
 Best Local Similarity 36.0%; Pred. No. 9.9e-09;
 Matches 45; Conservative 20; Mismatches 51; Indels 9; Gaps 3;
 QY 28 DVHLGLAMLVGSMCLFIAGFANGPPIPLMLSEIFPLHIKVGATGCVLTNMFAPLV 87
 DB 357 DYEAFSEFCIVAILLYVAFFELGPGIPMFIVAELESOGPPRAVAIVAGCCMTSNFLV 416
 QY 88 TKEFNSIMELTRPGAFWLTAAFCILSVLTFTLVPEETKGRTEOIT-----AHLRDD 141
 DB 417 GMLFSPAAYLVAY-VFTIFAAFLIFLITFFKVPETKGRTEEDTARAFEGAAH--SGK 473
 QY 142 GPLSV 146
 DB 474 GPAGV 478

RESULT 15
 GTR2_HUMAN STANDARD; PRT; 524 AA.
 ID GTR2_HUMAN
 AC P1168;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE GLUCOSE TRANSPORTER TYPE 2, LIVER.
 GN SLC2A2 OR GLUT2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OC NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RX TISSUE=Liver, and Kidney;
 RX MEDLINE=88289735; PubMed=3399500;
 RA Fukumoto H., Seino S., Imura H., Sieno Y., Eddy R.L., Fukushima Y.,
 RA Byers M.G., Shows T.B., Bell G.I.;
 RA "Sequence, tissue distribution, and chromosomal localization of mRNA
 RA encoding a human glucose transporter-like protein.";
 RT Proc. Natl. Acad. Sci. U.S.A. 85:5434-5438(1988).
 CC -!- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOMER LIKELY
 CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
 CC MEMBRANE OF HEPATOCYTES & IS RESPONSIBLE FOR UPTAKE OF GLUCOSE BY

CC THE BETA CELLS; IT MAY COMPRISE PART OF THE GLUCOSE-SENSING
 CC MECHANISM OF THE BETA CELL. IT MAY ALSO PARTICIPATE WITH THE
 CC NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCELLULAR TRANSPORT OF
 CC GLUCOSE IN THE SMALL INTESTINE & KIDNEY.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: LIVER, INSULIN-PRODUCING BETA CELL, SMALL
 CC INTESTINE, AND KIDNEY.
 CC -!- SIMILARITY: HIGH TO OTHER GLUCOSE TRANSPORTERS.
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: J03810; AA59514.1; -;
 DR PIR: A31318; A31318.
 DR MIM: 138160; -;
 DR Interpro: IPR003663; Sugar_transport.
 DR Interpro: IPR003662; sub_transport.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGTRANSPORT.
 DR PRINTS: PR00172; GLUCTRSPORT.
 DR PRINTS: PR01191; GLUCTRSPORT2.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 DR Transmembrane: Sugar transport; Transport; Glycoprotein;
 KM Multigene family; Polymorphism.
 FT TRANSMEM 11 31 POTENTIAL.
 FT DOMAIN 32 98 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 99 119 POTENTIAL.
 FT TRANSMEM 128 148 POTENTIAL.
 FT TRANSMEM 159 179 POTENTIAL.
 FT TRANSMEM 188 208 POTENTIAL.
 FT TRANSMEM 218 238 POTENTIAL.
 FT DOMAIN 239 303 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 304 324 POTENTIAL.
 FT TRANSMEM 339 359 POTENTIAL.
 FT TRANSMEM 370 390 POTENTIAL.
 FT TRANSMEM 401 421 POTENTIAL.
 FT TRANSMEM 434 454 POTENTIAL.
 FT TRANSMEM 462 482 POTENTIAL.
 FT DOMAIN 483 524 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 68 68 P->L.
 FT VARIANT 68 68
 SQ SEQUENCE 524 AA; 57489 MW; DA600577207EC083 CRC64;

Query Match 20.1%; Score 179; DB 1; Length 524;
 Best Local Similarity 34.8%; Pred. No. 1.7e-08;
 Matches 39; Conservative 20; Mismatches 49; Indels 4; Gaps 2;
 QY 34 AMLAVGSMC---LFTAGFVANGPPIPLMLSEIFPLHIKVGATGCVLTNMFAPLVTK 90
 DB 394 SMASVYSMAITLFLVFFELGPGIPMFIVAELESOGPPRAALATAASNMWNCNIVL 453
 QY 91 FNSIMEILTRPGAFWLTAAFCILSVLTFTLVPEETKGRTEOITAHLRDGG 142
 DB 454 PQYADFCGPY-VFPLFAGVLLAFILTFEKKVPETKGRTEEDTARAFEGAAH--SGK 504

Search completed: February 13, 2002, 22:06:05
 Job time: 861 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 22:04:42 ; Search time 136.8 Seconds
(without alignments)
176.425 Million cell updates/sec

Title: US-09-516-493-10
Perfect score: 889
Sequence: 1 WRPLZNZMIPRAGIRHELV.....VTGSPELSLWRKSGVEWPMQ 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: SPTREMBL_17:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	627	70.5	478	11	Q9JZJ1
2	627	70.5	479	11	Q9JZJ1
3	590	66.4	477	11	Q9JZJ0
4	585	65.8	477	11	Q9JZJ0
5	585	65.8	477	11	Q9JZJ0
6	548	61.6	477	4	Q9NSC4
7	540	60.7	477	4	Q9NY64
8	282	31.7	507	4	Q9UCQ3
9	278	31.3	433	5	Q9V610
10	271	30.5	489	5	Q9V609
11	268	30.1	463	10	Q9MAA4
12	262	29.5	539	5	Q9VU17
13	257	28.9	348	10	Q9ZU87
14	248.5	28.0	465	5	Q9V848
15	239	26.9	444	5	Q9V3G0
16	232.5	26.2	621	6	Q9BE72
17	225	25.3	487	10	Q9FRJ3
18	218.5	24.6	515	10	Q9LNU8
19	213.5	24.0	490	10	Q9J916

20	212	23.8	457	2	P96742	P96742 bacillus su
21	209.5	23.6	483	10	Q9ZVM0	Q9ZVM0 arabidopsis
22	208.5	23.5	510	10	Q65322	Q65322 petunia hyb
23	207	23.3	477	10	Q9SCW9	Q9SCW9 arabidopsis
24	206.5	23.2	501	10	Q9XHW3	Q9XHW3 oryza sativ
25	204.5	23.0	516	10	Q04078	Q04078 vicia faba
26	202	22.7	521	10	Q02848	Q02848 arabidopsis
27	201	22.6	558	10	Q9TF2	Q9TF2 arabidopsis
28	199.5	22.4	491	5	Q9VNDJ	Q9VNDJ drosophila
29	198.5	22.3	522	10	Q9SXB1	Q9SXB1 arabidopsis
30	197	22.2	581	10	Q9LKH1	Q9LKH1 mesembryant
31	195	21.9	523	10	Q06312	Q06312 nicotiana t
32	194.5	21.9	378	10	Q00400	Q00400 arabidopsis
33	194.5	21.9	492	5	Q44826	Q44826 caenorhabdi
34	194.5	21.9	510	5	Q44827	Q44827 caenorhabdi
35	194	21.8	493	10	Q9FYG3	Q9FYG3 arabidopsis
36	193.5	21.8	461	10	Q9M100	Q9M100 arabidopsis
37	193.5	21.8	582	10	Q23492	Q23492 arabidopsis
38	193	21.7	474	10	Q04036	Q04036 arabidopsis
39	193	21.7	496	10	Q65799	Q65799 arabidopsis
40	193	21.7	734	10	Q9SYQ3	Q9SYQ3 arabidopsis
41	193	21.7	734	10	Q9LM67	Q9LM67 arabidopsis
42	192.5	21.7	339	10	Q9SCW6	Q9SCW6 arabidopsis
43	192.5	21.7	519	10	Q9ZS76	Q9ZS76 vitis vinif
44	192.5	21.7	580	10	Q9C757	Q9C757 arabidopsis
45	191.5	21.5	497	5	Q9VND1	Q9VND1 drosophila

ALIGNMENTS

```
RESULT 1
09JZJ1 PRELIMINARY: PRT: 478 AA.
AC 09JZJ1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE GLUCOSE TRANSPORTER.
GN GLUTX1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20138191; PubMed=10671487;
RA Ibbertson M.R., Uldry M.A., Thorens B.;
RT "GLUTX1: A novel mammalian glucose transporter expressed in the
RT central nervous system and insulin-sensitive tissues.";
RL J. Biol. Chem. 275:4607-4612(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL: AJ245935; CAB75729.1; -.
DR InterPro: IPR003662; sub_transporter.
DR InterPro: IPR003663; Sugar_transporter.
DR Pfam: PF00083; sugar_tr. 1.
DR PRINTS: PR00171; SUGRTNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane.
SO SEQUENCE 478 AA; 51458 MW; 95841FC1F18C9EE9 CRC64;
```

Query Match 70.5%; Score 627; DB 11; Length 478;
Best local Similarity 100.0%; Pred. No. 1.8e-50;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LVPISEAPADVHGLAMLAWSKCLPIAGFVCGPIPWILMSFIPPLHIGKVGATGCVL 78
DB 357 LVPISEAPADVHGLAMLAWSKCLPIAGFVCGPIPWILMSFIPPLHIGKVGATGCVL 416
QY 79 TNMFMAFLVTKFENSIIMEILRPYGAFWLTAAFCILSVLFTLTVPETKGRTEQITAH 136

Db 417 TNNFMAFLVTKERNIMEILRPYGAFWLTAAFCILSVLFTLVPEPKRTLEQITAH 474

RESULT 2

Q9JMA6

PRELIMINARY:

PRT: 479 AA.

AC Q9JMA6: 09JUN20: 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE GLUCOSE TRANSPORTER 8.
 GN GLUT8.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ishibashi K.;
 RT "Molecular cloning of a new putative glucose transporter";
 RL Submitted (Oct-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL; AB033418; BAA94383.1; -;
 DR InterPro: IPR003662; sub_transporter.
 DR InterPro: IPR003663; Sugar transporter.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGRTNSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane.
 SQ SEQUENCE 479 AA; 51601 MW; EAB2B67B8CDD2PCF CRC64;

Query Match

Best Local Similarity 70.5%; Score 627; DB 11; Length 479;
 Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LVPISEAPADVHLGLAMLVAGSMCLFIAGFAVGMPILPMLNSEIFPLHIKGVATGCVL 78
 DB 358 LVPISEAPADVHLGLAMLVAGSMCLFIAGFAVGMPILPMLNSEIFPLHIKGVATGCVL 417
 OY 79 TNNFMAFLVTKERNIMEILRPYGAFWLTAAFCILSVLFTLVPEPKRTLEQITAH 136
 DB 418 TNNFMAFLVTKERNIMEILRPYGAFWLTAAFCILSVLFTLVPEPKRTLEQITAH 475

RESULT 3

Q9JUN20

PRELIMINARY:

PRT: 477 AA.

AC Q9JUN20: 09JUN20: 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE GLUCOSE TRANSPORTER.
 GN GLUTX1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20138191; PubMed=10671487;
 RA Ibberson M.R., Uldry M.A., Thorens B.;
 RT "GLUTX1: A novel mammalian glucose transporter expressed in the
 central nervous system and insulin-sensitive tissues";
 RL J. Biol. Chem. 275:4607-4612(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL; AJ245936; CAB75719.1; -;
 DR MGD; MGI:1860103; GLUTX1.
 DR InterPro: IPR003662; sub_transporter.

DR InterPro: IPR003663; Sugar_transporter.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGRTNSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane.
 SQ SEQUENCE 477 AA; 51578 MW; 14FF23DB9060DAF1 CRC64;

Query Match

Best Local Similarity 66.4%; Score 590; DB 11; Length 477;
 Matches 107; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

OY 19 LVPISEAPADVHLGLAMLVAGSMCLFIAGFAVGMPILPMLNSEIFPLHIKGVATGCVL 78
 DB 356 LVPISEAPADVHLGLAMLVAGSMCLFIAGFAVGMPILPMLNSEIFPLHIKGVATGCVL 415
 OY 79 TNNFMAFLVTKERNIMEILRPYGAFWLTAAFCILSVLFTLVPEPKRTLEQITAH 136
 DB 416 TNNFMAFLVTKERNIMEILRPYGAFWLTAAFCILSVLFTLVPEPKRTLEQITAH 473

RESULT 4

Q9JUP4

PRELIMINARY:

PRT: 477 AA.

AC Q9JUP4: 09JUN20: 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE GLUCOSE TRANSPORTER 8.
 GN GLUTX1 OR GLUT8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=TESTIS;
 RX MEDLINE=20283667; PubMed=10821868;
 RA Doege H., Schermann A., Bahrenberg C., Brauers A., Joost H.G.;
 RT "Glut8: A novel member of the sugar transport facilitator family with
 glucose transport activity";
 RL J. Biol. Chem. 275:16275-16280(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL; Y17802; CAB89815.1; -;
 DR MGD; MGI:1860103; GLUTX1.
 DR InterPro: IPR003662; sub_transporter.
 DR InterPro: IPR003663; Sugar_transporter.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGRTNSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane.
 SQ SEQUENCE 477 AA; 51507 MW; 59A985B52ADF478 CRC64;

Query Match

Best Local Similarity 65.8%; Score 585; DB 11; Length 477;
 Matches 106; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

OY 19 LVPISEAPADVHLGLAMLVAGSMCLFIAGFAVGMPILPMLNSEIFPLHIKGVATGCVL 78
 DB 356 LVPISEAPADVHLGLAMLVAGSMCLFIAGFAVGMPILPMLNSEIFPLHIKGVATGCVL 415
 OY 79 TNNFMAFLVTKERNIMEILRPYGAFWLTAAFCILSVLFTLVPEPKRTLEQITAH 136
 DB 416 TNNFMAFLVTKERNIMEILRPYGAFWLTAAFCILSVLFTLVPEPKRTLEQITAH 473

RESULT 5

Q9JIF3

PRELIMINARY:

PRT: 477 AA.

```

AC 09JIF3;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE GLUCOSE TRANSPORTER GLUT8.
GN GLUT8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RA MEDLINE=20319023; Pubmed=10860996;
RA Carayannopoulos M.O., Chi M.M.Y., Cui Y., Pingsterhaus J.M.,
RA McKnight R.A., Mueckler M., Devaskar S.U., Moley K.H.;
RT "Glut8 is a glucose transporter responsible for insulin-stimulated
RT glucose uptake in the blastocyst."
RL Proc. Natl. Acad. Sci. U.S.A. 97:7313-7318(2000).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
EMBL: AF232061; AAF78366.1; -.
DR MGD; MGI:1860103; Glutx1.
DR InterPro: IPR003662; sub_transporter.
DR InterPro: IPR003663; Sugar_transporter.
DR Pfam: PF00083; sugar_tr; 1.
DR PRINTS: PR00171; SUGRTNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KM Transmembrane.
SQ SEQUENCE 477 AA: 51523 MW: A375FB34E452F9A CRC64:

Query Match 65.8%; Score 585; DB 11; Length 477;
Best Local Similarity 89.8%; Pred. No. 1,4e-46;
Matches 106; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 19 LVPSAPADVHLGLAMLA VGSMLFTAGFAVNGCPILPMLMSEIFPHIRGVATGVLTN 78
DB 356 LVPIAAPRVQVGLAMLA VGSMLFTAGFAVNGCPILPMLMSEIFPLHVGVATGICVLN 415
QY 79 TNMFMAFLVTKFESSIMEILRPYGAFWLTAFCILSVLFTLPVPEYKGRILEQITAH 136
DB 416 TNMFMAFLVTKFESSVMEILRPYGAFWLTAFCILSVLFTLPVPEYKGRILEQITAH 473

RESULT 6
Q9NSC4 PRELIMINARY; PRT; 477 AA.
AC Q9NSC4;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE GLUCOSE TRANSPORTER 8.
GN GLUT8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA MEDLINE=20283667; Pubmed=10821868;
RA Doege H., Schuermann A., Bahrenberg C., Brauers A., Joost H.G.;
RT "Glut8: A novel member of the sugar transport facilitator family with
RT glucose transport activity."
RL J. Biol. Chem. 275:16275-16280(2000).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
EMBL: Y17801; CAB89809.1; -.
DR InterPro: IPR003662; sub_transporter.
DR InterPro: IPR003663; Sugar_transporter.
DR Pfam: PF00083; sugar_tr; 1.

```

```

DR PRINTS: PR00171; SUGRTNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KM Transmembrane.
SQ SEQUENCE 477 AA: 50792 MW: 0B480F94B40AEE76 CRC64:

Query Match 61.6%; Score 548; DB 4; Length 477;
Best Local Similarity 83.6%; Pred. No. 4e-43;
Matches 97; Conservative 11; Mismatches 8; Indels 0; Gaps 0;

QY 21 PLSAPADVHLGLAMLA VGSMLFTAGFAVNGCPILPMLMSEIFPHIRGVATGVLTN 80
DB 358 PVSAPVDASVGLAMLA VGSMLFTAGFAVNGCPILPMLMSEIFPLHVGVATGICVLN 417
QY 81 WMFAFLVTKFESSIMEILRPYGAFWLTAFCILSVLFTLPVPEYKGRILEQITAH 136
DB 418 WMFAFLVTKFESSLMEILRPYGAFWLTAFCILSVLFTLPVPEYKGRILEQITAH 473

RESULT 7
Q9NY64 PRELIMINARY; PRT; 477 AA.
AC Q9NY64;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE GLUCOSE TRANSPORTER.
GN GLUTX1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20138191; Pubmed=10671487;
RA Idberson M.R., Uldry M.A., Thorens B.;
RT "GLUTX1: A novel mammalian glucose transporter expressed in the
RT central nervous system and insulin-sensitive tissues."
RL J. Biol. Chem. 275:4607-4612(2000).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
EMBL: AJ245937; CAB75702.1; -.
DR InterPro: IPR003662; sub_transporter.
DR InterPro: IPR003663; Sugar_transporter.
DR Pfam: PF00083; sugar_tr; 1.
DR PRINTS: PR00171; SUGRTNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KM Transmembrane.
SQ SEQUENCE 477 AA: 50858 MW: 39DA9E2DBFFD3E2 CRC64:

Query Match 60.7%; Score 540; DB 4; Length 477;
Best Local Similarity 82.8%; Pred. No. 2.2e-42;
Matches 96; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 21 PLSAPADVHLGLAMLA VGSMLFTAGFAVNGCPILPMLMSEIFPHIRGVATGVLTN 80
DB 358 PVSAPVDASVGLAMLA VGSMLFTAGFAVNGCPILPMLMSEIFPLHVGVATGICVLN 417
QY 81 WMFAFLVTKFESSIMEILRPYGAFWLTAFCILSVLFTLPVPEYKGRILEQITAH 136
DB 418 WMFAFLVTKFESSLMEILRPYGAFWLTAFCILSVLFTLPVPEYKGRILEQITAH 473

RESULT 8
Q9UGO3 PRELIMINARY; PRT; 507 AA.
AC Q9UGO3;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)

```

DE SUGAR TRANSPORTER (GLUCOSE TRANSPORTER 9).
 GN SLC2A6 OR GLUT9.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RC TISSUE=LYMPHOID;
 RA Young J.M., Woodward K.J., Aziz S., Burley M., Kwiatkowski D.J.,
 RT Povey S.;
 RT "Cloning of a sugar transporter gene, a β -glucose subunit like gene and
 RT three novel genes in human chromosome 9q34.";
 RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Doege H., Bocianski A., Joost H.G., Schermann A.;
 RT "Activity and genomic organization of human glucose transporter 9
 RT (GLUT9), a novel member of the family of sugar-transport facilitators
 RT predominantly expressed in brain and leucocytes.";
 RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL: AJ011372; CAB6155.1;
 DR EMBL: Y17803; CAB9696.1;
 DR InterPro: IPR003662; sub_transporter.
 DR InterPro: IPR003663; Sugar-transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 DR PRINTS: PR00171; SUGRTNSPRT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Sugar transporter; Transmembrane.
 SO SEQUENCE 507 AA; 54569 MW; 1FC62EB11588460 CRC64;
 Query Match 31.7%; Score 282; DB 4; Length 507;
 Best Local Similarity 44.3%; Pred. No. 2.6e-18;
 Matches 62; Conservative 22; Mismatches 52; Indels 4; Gaps 2;
 QY 2 RPLZNMIPRAAGIRHETLPIAEPADYHIC-LAMIAVGSCLFIAGFANGWGPIMILM 60
 DB 365 RPLSPN---STAGLEFSWGDIAOPLAAPAGYTLVPLLTMTLMTFIMGYAVGWGPIITWLM 421
 QY 61 SEIRPLIKVANGCVLTWMPAFVLTKEFNSIMELIRYGAFWLTAACIIISVLETLT 120
 DB 422 SEVPLRLARGVASGLCVLAWMLTAFVLTKSLFVSTFGVQVFFPFAALCVLSVFTGC 481
 QY 121 FVPEKGRTELEQITAHLRDG 140
 DB 482 CPEETKGRSLLEQIESFFRWG 501
 RESULT 9
 Q9V610 PRELIMINARY: PRT; 433 AA.
 ID Q9V610
 AC Q9V610
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE CG8234 PROTEIN.
 GN CG8234
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Branton R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler A.,
 RA Burks K.C., Busan D.A., Butler H., Cadieu E., Gentler A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies I.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doop L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Liao X., Lel Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J.R., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Smith T.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Turner R., Venter E., Wang A.H., Wang X.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.C., Wu D., Yang S., Yao Q.A.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL: AE003825; AAF58630.1;
 DR FLYBase: FBgn0033644; CG8234.
 DR InterPro: IPR003662; sub_transporter.
 DR InterPro: IPR003663; Sugar-transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 DR PRINTS: PR00171; SUGRTNSPRT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 2.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane.
 SO SEQUENCE 433 AA; 47357 MW; 63F0E13046A769DD CRC64;
 Query Match 31.3%; Score 278; DB 5; Length 433;
 Best Local Similarity 44.7%; Pred. No. 5.1e-18;
 Matches 55; Conservative 21; Mismatches 43; Indels 4; Gaps 3;
 QY 24 AEPADYHIC-LAMIAVGSCLFIAGFANGWGPIMILMSEIFPLIKVANGCVLTWMP 82
 DB 302 AHGPVSHDG-WPLTGFVYITLIFSLGFGPIRMGELIPAKIRPASPASTAFANWF 359
 QY 83 MAVLTKEFNSIMELIRYGAFWLTAACIIISVLETLTFVPEKGRTELEQITAHLRDGC 142
 DB 360 CTFVVTKTFQDLTYAMGAHGAFWLFGALICIVGLFFVITFVPEKGRSLLETERKMM-GRV 418
 QY 143 PLS 145
 DB 419 PMS 421
 RESULT 10
 Q9V609 PRELIMINARY: PRT; 489 AA.
 ID Q9V609
 AC Q9V609
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)

CG7801. PROTEIN.
Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227:
[1]
SEQUENCE FROM N.A.
STRAIN-BERKELEY:
MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Goodyne J.D.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champagne M., Pfeiffer B.D.,
Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abbil J.F., Ayhyan A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durdin K.J., Evangelista C.C., Ferraz C., Fertler S., Fleischmann W.,
Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbegam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Liao X., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Malteli B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Paclet J.M.,
Palazzo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Sidenkiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
Williams S.M., Woodage T., Wolley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zhang X.-H., Zhong F., Zhong W., Zhou X., Zhou S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
"The genome sequence of Drosophila melanogaster."
Science 287:2185-2195(2000).
-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
-1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
EMBL: AEO03825; AAF58631.1; -
DR FLYbase: FBgn0033643; CG7801.
DR InterPro: IPR003662; sub_transporter.
DR InterPro: IPR003663; Sugar_transporter.
DR Pfam: PF00083; sugar_tr_1.
DR PRINTS: PR00171; SUGTRNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane
SEQUENCE 489 AA; 53026 MW; D26436DAD69723C6 CRC64;

Query March	30.5%	Score 271	DB 5	Length 489
Best Local Similarity	46.2%	Pred. No. 2.6e-17		
Matches	48	Conservative	18	Mismatches 36
				Indels 2
				Gaps 1
QY	30	HLGLAMLVGSKCLPIAGFAVGWGPPIWLMSEIPPLHIKRGATGCVYLTNNFMFLYRK	89	
Db	364	HLG-MLPETCFVYIILGSLGFGPIPLMLMGELLPAKIRGSAASAVATFNMFCTFVVYRK	421	
QY	90	EPNSIMEILRPYGAWLTAFCILSVLFTLTFVPETKCTLEOI	133	
		:		

Db	422	TF0DLTVAMGAHGAFWLGCALCGFGLFVILIVETQSKTLEDI	465
RESULT	11		
Q9MAA4			
ID	Q9MAA4	PRELIMINARY;	PRT; 463 AA.
AC	Q9MAA4		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)		
DE	PURATIVE SUGAR TRANSPORTER.		
CN	T12H1.11.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eumariophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eupariophyta; Magnoliophyta; eudicotyledons; core eudicots; Brassicales; Brassicaceae; Arabidopsids.		
CC	NCBI_TaxID=3702;		
RM	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV, COLUMBIA;		
RA	Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,		
RA	Ronning C.M., Koo H., Fujii C.Y., Uterback T.R., Barnstead M.E.,		
RA	Bowman C.L., White O., Nierman W.C., Fraser C.M.;		
RT	"Arabidopsis thaliana chromosome II BAC T12H1 genomic sequence."		
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.		
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).		
CC	-1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.		
DR	EMBL: AC009177; AAF27021.1; -		
DR	InterPro: IPR003662; sub_Transporter.		
DR	InterPro: IPR003663; Sugar_transporter.		
DR	Pfam: PF00083; sugar_tr.1.		
DR	PRINTS: PS00171; SUGRTNSPORT.		
DR	PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.		
KW	Sugar transport; Transmembrane.		
SO	SEQUENCE 463 AA; 50148 MW; 82B93947AE79DAC0 CRC64;		

RESULT	12		
09VU17			
ID	09VU17	PRELIMINARY;	PRT: 539 AA.
AC	09VU17;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	CG10960 PROTEIN.		
CN	CG10960.		
OS	Drosophila melanogaster (fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxId=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY;		
RA	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celisiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champé M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtils K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostlin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,
 RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spletter E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskis R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissensbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhao M., Zhang G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA "The genome sequence of *Drosophila melanogaster*."
 RT "The genome sequence of *Drosophila melanogaster*."
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL: AEO03540; AAF49874.1;
 DR FLYBase: FBgn0036316; CG10960.
 DR InterPro: IPR003662; sub_transporter.
 DR InterPro: IPR003663; Sugar_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 DR PRINTS: PR00171; SUGARTRNSPORT.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 DR Transmembrane.
 KW SUGAR_TRANSPORT_2; 1.
 SO SEQUENCE 539 AA; 58413 MW; A2F4814DBF676BBD CRC64;

Query Match 29.5%; Score 262; DB 5; Length 539;
 Best Local Similarity 44.9%; Pred. No. 2e-16; Mismatches 38; Indels 0; Gaps 0;
 Matches 48; Conservative 21; Mismatches 38; Indels 0; Gaps 0;

QY 31 LGTAWLAVGSMCLFIAGFAVNGKPIPLMLMSLFPPLIKNGVATGCVLTWMPFAVLTKE 90
 DB 419 VSLGMLFVSLCLFIIMFISGYGVPVLMGMLGELFARDIKGFAAGSLAGTSMWLAFFVYTKT 478
 QY 91 ENSIMEILRPYGAFLTAACILSVLTFTFVETGRTLEQTTAHL 137
 DB 479 FVNLNGSLGIGTFWFLAGLTIVGVLFVFAVETKGSINLEQDEL 525

RESULT 13
 ID Q92087 PRELIMINARY; PRT; 348 AA.
 AC Q92087;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 GN PUTATIVE SUGAR TRANSPORTER.
 OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
 RA Barnstead M.E., Mason T.M., Bowman C.L., Rensing C.M., Benito M.,
 RA Carreira A.J., Creasy T.H., Bueli C.R., Town C.D., Nierman W.C.,
 RA Fraser C.M., Venter J.C.;
 RT "Arabidopsis thaliana chromosome II BAC T9J23 genomic sequence."
 RL Submitted (Feb-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL: AC006072; AAD13706.1;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
 DR Sugar transporter; Transmembrane.
 KW SUGAR_TRANSPORT_1; UNKNOWN_1.
 SO SEQUENCE 348 AA; 36899 MW; C9A7C75133FF69BF CRC64;

Query Match 28.9%; Score 257; DB 10; Length 348;
 Best Local Similarity 44.5%; Pred. No. 3.6e-16;
 Matches 53; Conservative 20; Mismatches 32; Indels 14; Gaps 2;

QY 17 HELVPISAEPAVDHLGLAWLAVGSMCLFIAGFAVNGKPIPLMLMSLFPPLIKNGVATGVC 76
 DB 240 HEAVP-----LAVGVIMYITGFSFGKGMKMPVMSSELPINIGVAGGMA 286
 QY 77 VLTWMPFAFLVTKFENSIMEILRPYGAFLTAACILSVLTFTFVETGRTLEQTTAHL 135
 DB 287 TLVWMPGAMVAVSTFNLMS-WSSYGTFLYLAIMALATVFAIVETKGTLEQTTAHL 344

RESULT 14
 ID Q9V848 PRELIMINARY; PRT; 465 AA.
 AC Q9V848;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CG6484 PROTEIN.
 GN CG6484.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champé M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtils K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostlin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mervlov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sidenkiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL: AE003802; AAF57829.1; -;
 DR FlyBase: FBgn0034247; CG6484.
 DR InterPro: IPR003662; sub_transporter.
 DR InterPro: IPR003663; Sugar_transporter.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGTRNSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 2.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane.
 SQ SEQUENCE 465 AA; 50271 MW; 004B4C4CF406BDBB CRC64;

Query Match 28.0%; Score 248.5; DB 5; Length 465;
 Best local Similarity 38.1%; Pred. No. 3.1e-15;
 Matches 45; Conservative 27; Mismatches 45; Indels 1; Gaps 1;
 QY 25 EPADYHLGLAVGSMCLFIAGVAGMGPIPWLMSEIFPLIKGVATGCVLTNNMFA 84
 Db 340 DPASMD-NFGMLPISICIFITFISIGCPVIMVAMELFSDEKSVASISAGTSMWLSA 398
 QY 85 FLVTKREFSIMEILRPYGAFWLTAFCILSVLFTLFVETVETGRTLEQITAHLRDGDG 142
 Db 399 FVYTLFLPILKSSIGCPFTFWIFTAIAVIAFYSLFVPEVETGKTIIEIDLLSGKG 456
 RESULT 15
 Q9V3G0 PRELIMINARY; PRT; 444 AA.
 AC Q9V3G0;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
 DE SUT4 PROTEIN
 GN SUT4 OR CG1380.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins K.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe C.R., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.M., Miklos G.L.G.,
 RA Abiri J.F., Agbayani A., An H.-J., Andrews-Piankoff C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Chiriac S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegum C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mervlov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sidenkiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Andersson Escher S., Lohlaht M., Rasmussen-Lestander A.;
 RT "Sugar transporters in Drosophila melanogaster."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL: AE003831; AAF58847.1; -;
 DR EMBL: AF199486; AAF13274.1; -;
 DR FlyBase: FBgn0028560; sut4.
 DR InterPro: IPR003662; sub_transporter.
 DR InterPro: IPR003663; Sugar_transporter.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGTRNSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Sugar transporter; Transmembrane.
 SQ SEQUENCE 444 AA; 48028 MW; D18068459A98661A CRC64;

Query Match 26.9%; Score 239; DB 5; Length 444;
 Best local Similarity 39.6%; Pred. No. 2.2e-14;
 Matches 42; Conservative 27; Mismatches 37; Indels 0; Gaps 0;
 QY 32 GLAVLAVGSMCLFIAGVAGMGPIPWLMSEIFPLIKGVATGCVLTNNMFAFLVTKRF 91
 Db 334 GYVGMILAVGCVFLIGSLFCGPVIMVAMELFAEDVKALAGSIAGTNNCFIVLLF 393
 QY 92 NSIMEILRPYGAFWLTAFCILSVLFTLFVETVETGRTLEQITAHLRDGDG 137
 Db 394 PVLDIIIGATACFAIFFGFAVAVFLFLIPETGKTLNIDIQKM 439

Search completed: February 13, 2002, 22:04:43
 Job time: 864 sec

Thu Feb 14 07:44:05 2002

us-09-516-493-10.rspt

Page 8

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 21:50:12 ; Search time 104.41 Seconds

(without alignments)
66,688 Million cell updates/sec

Title: US-09-516-493-12

Perfect score: 508
Sequence: 1 EPADVHGLAMWLAAGSMCLF.....RPYGFMTLTAACILSLVLT 94

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_1101.*
2: /SID52/gcgdata/geneseq/AA1980.DAT.*
3: /SID52/gcgdata/geneseq/AA1981.DAT.*
4: /SID52/gcgdata/geneseq/AA1982.DAT.*
5: /SID52/gcgdata/geneseq/AA1983.DAT.*
6: /SID52/gcgdata/geneseq/AA1984.DAT.*
7: /SID52/gcgdata/geneseq/AA1985.DAT.*
8: /SID52/gcgdata/geneseq/AA1986.DAT.*
9: /SID52/gcgdata/geneseq/AA1987.DAT.*
10: /SID52/gcgdata/geneseq/AA1988.DAT.*
11: /SID52/gcgdata/geneseq/AA1989.DAT.*
12: /SID52/gcgdata/geneseq/AA1990.DAT.*
13: /SID52/gcgdata/geneseq/AA1991.DAT.*
14: /SID52/gcgdata/geneseq/AA1992.DAT.*
15: /SID52/gcgdata/geneseq/AA1993.DAT.*
16: /SID52/gcgdata/geneseq/AA1994.DAT.*
17: /SID52/gcgdata/geneseq/AA1995.DAT.*
18: /SID52/gcgdata/geneseq/AA1996.DAT.*
19: /SID52/gcgdata/geneseq/AA1997.DAT.*
20: /SID52/gcgdata/geneseq/AA1998.DAT.*
21: /SID52/gcgdata/geneseq/AA1999.DAT.*
22: /SID52/gcgdata/geneseq/AA2000.DAT.*
23: /SID52/gcgdata/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	508	100.0	478	22	AAB66933
2	487	95.9	478	22	AAB66939
3	482	94.9	477	22	AAB66934
4	457	90.0	477	22	AAE04888
5	454	89.4	477	22	AAB66932
6	220	43.3	105	21	AAV95019
7	220	43.3	106	21	AAV95019
8	220	43.3	147	21	AAV95032
9	217	42.7	507	22	AAB66937
10	217	42.7	507	22	AAB66941
11	214	42.1	503	22	AAB66938

12	189	37.2	70	22	AAM25722	Human protein sequ
13	187.5	36.9	193	21	AA37400	Arabidopsis thalia
14	187.5	36.9	193	21	AA37422	Arabidopsis thalia
15	187.5	36.9	212	21	AA37398	Arabidopsis thalia
16	187.5	36.9	212	21	AA37420	Arabidopsis thalia
17	187.5	36.9	437	21	AA313750	Arabidopsis thalia
18	187.5	36.9	437	21	AA320978	Arabidopsis thalia
19	187.5	36.9	438	21	AA38871	Arabidopsis thalia
20	187.5	36.9	447	21	AA313749	Arabidopsis thalia
21	187.5	36.9	447	21	AA320977	Arabidopsis thalia
22	187.5	36.9	448	21	AA388870	Arabidopsis thalia
23	187.5	36.9	453	21	AA313748	Arabidopsis thalia
24	187.5	36.9	463	21	AA320976	Arabidopsis thalia
25	187.5	36.9	464	21	AA388869	Arabidopsis thalia
26	163.5	32.2	526	20	AAV02168	A facilitative glu
27	157.5	31.0	237	21	AA306143	Arabidopsis thalia
28	157.5	31.0	294	21	AA306142	Arabidopsis thalia
29	157.5	31.0	352	21	AA306141	Arabidopsis thalia
30	157.5	31.0	471	21	AA388866	Arabidopsis thalia
31	157.5	31.0	473	21	AA316800	Arabidopsis thalia
32	157.5	31.0	473	21	AA352317	Arabidopsis thalia
33	157.5	31.0	491	21	AA316799	Arabidopsis thalia
34	157.5	31.0	491	21	AA352316	Arabidopsis thalia
35	157.5	31.0	522	21	AA316798	Arabidopsis thalia
36	157.5	31.0	522	21	AA352315	Arabidopsis thalia
37	157.5	31.0	558	21	AA388865	Arabidopsis thalia
38	157.5	31.0	564	21	AA388864	Arabidopsis thalia
39	155.5	30.6	502	20	AA496625	Corn hexose carie
40	154.5	30.4	420	21	AA330434	Arabidopsis thalia
41	154.5	30.4	477	21	AA330433	Arabidopsis thalia
42	154.5	30.4	483	21	AA330432	Arabidopsis thalia
43	154.5	30.4	618	22	AAB66936	Rat GLUTX2. Ratu
44	151	29.7	508	21	AA329530	Arabidopsis thalia
45	151	29.7	555	21	AA329529	Arabidopsis thalia

ALIGNMENTS

RESULT 1	
ID AAB66933	standard; Protein: 478 AA.
AC AAB66933:	
DT 17-APR-2001	(first entry)
DE Rat GLUTX1.	
XX	
XX	
XX	
KW Rat: GLUTX: gene therapy; vaccine; hexose transport modulator;	
KW hexose transport disorder; ischemia; diabetes; hyperglycemia;	
KW hypoglycemia; glucose metabolism disorder; neurodegenerative disease.	
XX	
OS Rattus sp.	
PN	
XX	
XX	
PD 18-JAN-2001.	
XX	
XX	
PF 14-JUL-2000; 2000MO-IB01042.	
XX	
PR 14-JUL-1999; 99US-0143907.	
PR 27-AUG-1999; 99US-0151140.	
PR 23-FEB-2000; 2000US-0184285.	
PR 13-JUL-2000; 2000US-0616132.	
XX	
PA (UYLA-) UNIV LAUSANNE.	
XX	
PI Thorens B, Ibberson M, Uldry M.	
XX	
DR WPI: 2001-112615/12.	
XX	
DR N-PSDB: AAF55866.	
XX	

PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -
 XX
 PS Claim 11; Page 71-72; 124pp; English.

XX The present invention relates to GLUTX proteins (AA55865-AA55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is rat GLUTX1.
 XX

Sequence 478 AA:

Query Match 100.0%; Score 508; DB 22; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1.3e-55;
 Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EPADVHLGLAMLVGSMCLFTAGFAVNGRPFWLMSLFFPLHKGATGVCVLTNMFMA 60
 Db 363 epadvhlglamlvagsmclftagfavngpfpwlmselfplhkyatgvcvltnwima 422

OY 61 FLVTKFNSIMEILRPYGAFWLTAFCILSVLFT 94
 Db 423 flvtkfnsimeilrpygafwlttafcilsvlft 456

RESULT 2

AAB66939 ID AAB66939 standard; Protein; 478 AA.

XX AAB66939;

XX 17-APR-2001 (first entry)

XX GLUTX1 consensus sequence.

XX GLUTX; gene therapy; vaccine; hexose transport modulator; human; rat;
 KM hexose transport disorder; ischaemia; diabetes; hyperglycaemia; murine;
 KM hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX

OS Homo sapiens.

OS Mus sp.

OS Rattus sp.

PN W0200104145-A2.

PD 18-JAN-2001.

PF 14-JUL-2000; 2000MO-IB01042.

PR 14-JUL-1999; 99US-0143907.

PR 27-AUG-1999; 99US-0151140.

PR 23-FEB-2000; 2000US-0184285.

PR 13-JUL-2000; 2000US-0616132.

PA (UYLA-) UNIV LAUSANNE.

PI Thorens B, Idberson M, Uldry M;

DR WPI; 2001-112615/12.

XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -
 XX
 PS Claim 11; Page 74-75; 124pp; English.

XX The present invention relates to GLUTX proteins (AA55865-AA55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is rat GLUTX1.
 XX

CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is a consensus sequence
 CC for GLUTX1.
 XX

Sequence 478 AA:

Query Match 95.9%; Score 487; DB 22; Length 478;
 Best Local Similarity 92.6%; Pred. No. 5.1e-53;
 Matches 87; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 EPADVHLGLAMLVGSMCLFTAGFAVNGRPFWLMSLFFPLHKGATGVCVLTNMFMA 60
 Db 363 epadvsvglamlvagsmclftagfavngpfpwlmselfplhkyatgvcvltnwima 422

OY 61 FLVTKFNSIMEILRPYGAFWLTAFCILSVLFT 94
 Db 423 flvtkfnsimeilrpygafwlttafcilsvlft 456

RESULT 3

AAB66934 ID AAB66934 standard; Protein; 477 AA.

XX AAB66934;

XX 17-APR-2001 (first entry)

XX Murine GLUTX1.

XX Murine; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KM hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
 KM hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX

OS Mus sp.

PN W0200104145-A2.

PD 18-JAN-2001.

PF 14-JUL-2000; 2000MO-IB01042.

PR 14-JUL-1999; 99US-0143907.

PR 27-AUG-1999; 99US-0151140.

PR 23-FEB-2000; 2000US-0184285.

PR 13-JUL-2000; 2000US-0616132.

PA (UYLA-) UNIV LAUSANNE.

PI Thorens B, Idberson M, Uldry M;

DR WPI; 2001-112615/12.

DR N-PDB; AA55867.

XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -
 XX
 PS Claim 11; Page 73-74; 124pp; English.

XX The present invention relates to GLUTX proteins (AA55865-AA55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is murine GLUTX1.
 XX

Sequence 477 AA:

Query Match	94.9%	Score 482;	DB 22;	Length 477;
Best Local Similarity	91.5%;	Pred. No. 2.4e-52;		
Matches	86;	Conservative	5;	Mismatches 3;
				Indels 0;
				Gaps 0;

Oy 1 EPADVHLGLAWLAVGSMCLFIAGFAVCWGRIPIMLIMSEIFRLHIKGVATGCVLTNMMA 60
 ||| : ||||||||||||||||||||||||||||| : ||||| : |||||
Db 362 epvddvqvgqlawlavgsmlcfiaqfavgwqpripwlmselrplhvqvatgclvcltnmfa 422

```

QY      61 FLVTKEFNSIMEILRPYGAFWLTAFCILSVLFT 94
         |||||:|||||||:|||||||:|||||||:|||||||
Db      422 flvtkefnsvmemlrrpygaflwtaafcalsvlft 45

```

RESULT	4
AAE04888	
ID	AAE04888 standard; Protein: 477 AA.

AC	AAE04888;
XX	
DT	10-SEP-2001 (first entry)

Human transporter and ion channel-1 (TRICH-1) protein.

Human, transporter and ion channel-1, TRICH-1; vaccine; cystic fibrosis; gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy; hypertension; angina; neurological disorder; asthma; bipolar disorder; dementia; depression; Alzheimer's disease; epilepsy; mood; arrhythmia; Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke; Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy; demyelinating disease; mental disorder; Schizophrenia; polypositis; muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease; dermatomyositis; diabetes mellitus; immunological disorder; psoriasis; rheumatoid arthritis; Sjögren's syndrome; systemic lupus erythematosus; sickle cell anaemia; Wilson's disease; infertility; Cushing's disease; scleroderma; pulmonary artery stenosis; noctopic; Addison's disease; malabsorption syndrome; hypercholesterolaemia; cancer.

OS Homo sapiens.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FT	Domain	25..474	/note= "Sugar transporter domain"
FT	Domain	259..279	

F1	/label= lalissmembrane_dollall
F1T	Domain 293..313

FT	Domain	320..339
F1	/label-11ansimembrade_domain1	

FT	Domain	
F1		/label=11dismembrane_domain
		438..457

```

F1
XX
/label= transmembrane_domain

```

MOZ0014000 AZ
XX
XX

FD-302a (Rev. 11-27-2011)

22 DEC 2000, 2000MC-0332033.
XX

PR 14-JAN-2000: 2000US-0176083.

PR 28-JAN-2000; 2000US-0178572.

PR 10-FEB-2000; 2000US-0181625.

PA (INCY-) INCYTE GENOMICS INC.

Baughn MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy K, Lal P;

PI Tang YT, Khan FA;

WPI; 2001-418042/44.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

XX Novel human transporter and ion channel proteins useful for treating
PT and preventing transport, neurological, muscle and immunological
PM disorders -
XX
PS Claim 1; Page 112-113, 160pp; English.

PS Claim 1; Page 112-113; 160pp; English.

CC The present sequence is transport and ion channel-1 (TRICH-1) protein.
CC TRICH is used as vaccine. TRICH is useful for treating a disease or
CC condition associated with decreased expression of functional TRICH,
CC such as transport disorder including amyotrophic lateral sclerosis,
CC cystic fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth
CC disease, Duchenne muscular dystrophy, angina and hypertension.
CC neurological disorders including Alzheimer's disease, amnesia, bipolar
CC disorder, dementia, depression, epilepsy, ischemic cerebrovascular
CC disease, stroke, cerebral neoplasms, Pick's disease, Huntington's
CC disease and Parkinson's disease, demyelinating diseases, mental disorders
CC including mood, anxiety, schizophrenia and seasonal affective disorder,
CC muscle disorder including cardiomyopathy, myocarditis, polymyositis,
CC dermatomyositis, arrhythmias and asthma and immunological disorders
CC including AIDS, adult respiratory distress syndrome (ARDS), allergies,
CC anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
CC syndrome, systemic lupus erythematosus and other diseases including
CC sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary
CC artery stenosis, Grave's disease, Cushing's disease, Addison's disease,
CC glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers
CC porphyras and viral, bacterial, fungal, helminthic and protozoal
CC infections. TRICH DNA is useful in gene therapy and in diagnostic
CC purposes.

AA	Sequence	477	AA;
50			

Query Match	90.0%	Score 457;	DB 22;	length 477;
Best Local Similarity	85.1%	Pred. No. 3.3e-49;		
Matches 80; Conservative	8;	Mismatches 6;	Indels 0;	Gaps 0;

```
QY      1 EPADVHLGLAMLAVGSNCLFIAGFAVGWSPITPWLIMSEIFPLHIKGVATGVCVLTNWFMA 60
      : 1 : :::::::::::::::::::::::::::::::::::::::::::::::::::: 11
Db      362 qpvdasvlgwlaavgsncifgiqfavgwspipwlmselfplhvkvatgicvltlnwlna 422
```

```
Qy      61 FLVTKEFNSIMEILRPYGAWLTTAAFCILSVLT 94
        |||||:::|||||::||| ||||
Db     422 flvtkefsslmelrpygafwlasafcifsvlt 455
```

RESULT 5

AAB00552
 ID AAB66932 standard; Protein; 477 AA.

AA
AC AAB66932;

DT 17-APR-2001 (first entry)

Human GLUTX1.

Human; GLUTX; KW

hypoglycaemia

OS Homo sapiens.

PN WO200104145-A2

PD 18-JAN-2001.

PF 14-JUL-2000; 2000WO-1B01042.

PR 14-JUL-1999; 99US-0143907.

PR 23-FEB-2000; 2000US-0184285.

	X	:
	C	:C
	M	:M
	R	:R
	G	:G
/	L	:L
B	S	:S
<	O	:O
>	H	:H
<	C	:C
>	F	:F
E	T	:T

PA (UYLA-) UNTIV LAUSANNE.
 XX
 PI Thorens B, Ibberson M, Uidry M;
 XX WPI: 2001-112615/12.
 DR N-PSDB: AAF53865.
 XX
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. Ischemia and diabetes -
 XX
 PS Claim 11; Page 70-71; 124pp: English.
 XX
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as Ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is human GLUTX1.
 XX
 SQ Sequence 477 AA:
 Query Match 89.4%; Score 454; DB 22; Length 477;
 Best Local Similarity 84.0%; Pred. No. 7.9e-49;
 Matches 79; Conservative 9; Mismatches 6; Indels 0; Gaps 0;
 QY 1 EPADVHLGLAVLAVGSMCLFIFGAVNGKPIPWMLMSEIFPLHIKGVATGCVLTNMFMA 60
 DB 362 qpvdasvglawlavaygmclfiagfavwqpiplmwseilphvayalcvtlnwlna 421
 QY 61 PLVTKEFNSEIMEILRPYGAFLMTAAFCILSVLT 94
 DB 422 flvtkefseimeilrpygaflwlasatcitsvlft 455
 RESULT 6
 AAY95019
 ID AAY95019 standard; Protein: 105 AA.
 XX
 AC AAY95019;
 XX
 DT 19-JUN-2000 (first entry)
 XX
 DE Human secreted protein vq1_L, SEQ ID NO:78.
 XX
 KW Human; secreted protein; cancer; tumour; cardiovascular disorder;
 KW blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;
 KW infection; fungal; bacterial; viral; HIV; allergy; arthritis;
 KW neurodegenerative disease; asthma; contraceptive.
 XX
 OS Homo sapiens.
 XX
 PN WO200011015-A1.
 XX
 PD 02-MAR-2000.
 XX
 PF 24-AUG-1999; 99WO-US19351.
 XX
 PR 24-AUG-1998; 98US-0097638.
 PR 24-AUG-1998; 98US-0097659.
 PR 09-SEP-1998; 98US-0099618.
 PR 28-SEP-1998; 98US-0102092.
 PR 25-NOV-1998; 98US-0109978.
 PR 23-DEC-1998; 98US-0113645.
 PR 23-DEC-1998; 98US-0113646.
 PR 23-AUG-1999; 99US-0379246.
 XX
 PA (ALPH-) ALPHAGENE INC.
 XX
 PI Valenzuela D, Yuan O, Hoffman H, Hall J, Raplejo P;
 XX

DR WPI: 2000-224657/19.
 XX
 PI New secreted or transmembrane proteins and polynucleotides encoding
 PT them, useful for treating neurodegenerative disorders, autoimmune
 PT diseases and cancer -
 XX
 PS Claim 87; Page 335-336; 357pp: English.
 XX
 CC The invention relates to 40 human secreted proteins (AAY94981-Y95020),
 CC and cDNA sequences encoding them (AAB23423-A23462). The secreted
 CC proteins of the invention include those that are thought to be only
 CC partially secreted, i.e., transmembrane proteins. The proteins of the
 CC invention may exhibit one or more activities selected from the following:
 CC cytokine activity; cell proliferation; differentiation; immune
 CC modulation; haematopoiesis regulation; tissue growth activity;
 CC activin/inhibin activity; chemotactic/chemokinetic activity; hemostatic
 CC and thrombolytic activity; anti-inflammatory activity; and tumour
 CC inhibition activity. The proteins may be administered to patients as
 CC vaccines, and the nucleotides may be used as part of a gene therapy
 CC regime. Diseases or conditions that may be treated using the proteins or
 CC nucleotides of the invention include autoimmune diseases; genetic
 CC disorders; haemophilia; cardiovascular diseases; cancer; bacterial,
 CC fungal and viral infections, especially HIV; multiple sclerosis;
 CC rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome;
 CC insulin dependent diabetes mellitus; and allergic reactions such as
 CC asthma and anaemia. They may also be used for treating wounds, burns,
 CC ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease and amyotrophic
 CC lateral sclerosis (ALS). Proteins with activin/inhibin activity may
 CC additionally be useful as contraceptives. Nucleic acid sequences of the
 CC invention may be used in chromosome mapping, and as a source of
 CC diagnostic primers and probes. The present sequence represents one of the
 CC 40 proteins of the invention.
 XX
 SQ Sequence 105 AA:
 Query Match 43.3%; Score 220; DB 21; Length 105;
 Best Local Similarity 51.3%; Pred. No. 3.6e-20;
 Matches 39; Conservative 15; Mismatches 22; Indels 0; Gaps 0;
 QY 19 LEIAGFVAVGKPIPWMLMSEIFPLHIKGVATGCVLTNMFMAFLVTKEFNSEIMEILRPG 78
 DB 2 flimgyavwqpiplmwseilphvayalcvtlnwlna 421
 QY 79 AFWLTAFCILSVLT 94
 DB 62 plfftaalcivslft 77
 RESULT 7
 AAB45158
 ID AAB45158 standard; Protein: 106 AA.
 XX
 AC AAB45158;
 XX
 DT 12-FEB-2001 (first entry)
 XX
 DE Human secreted protein sequence encoded by gene 39 SEQ ID NO:99.
 XX
 KW Human; secreted protein; immunosuppressive; antiarthritic; anti-rheumatic;
 KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
 KW neurotropic; neuroprotective; antibacterial; vitucide; fungicide; cancer;
 KW ophthalmological; autoimmune disease; hyperproliferative disorder;
 KW cardiovascular disorder; cerebrovascular disorder; wound healing;
 KW nervous system disorder; aging; chemotaxis.
 XX
 OS Homo sapiens.
 XX
 PN WO200058467-A1.
 XX
 PD 05-OCT-2000.
 XX

PF		22-MAR-2000:	2000OWO-US07505.	
PR		26-MAR-1999:	99US-0126502.	
PR		17-DEC-1999:	99US-0172410.	
XX				
PA	(HUMA-)	HUMAN GENOME SCI INC.		
XX				
PT	Rosen CA, Ruben SM, Komatsoulis G;			
DR	WPI: 2000-611712/58.			
DR	N-PsDB: AAC80569.			
XX				
PT	Nucleic acid molecules encoding human secreted proteins, used in preventing, treating or ameliorating a disorder, e.g. Alzheimer's and Parkinson's diseases and cancers -			
PS	Claim 11; Page 396; 440pp: English.			
XX				
CC	polynucleotide sequences AACG0531-C80580 represent cDNA encoding human secreted proteins AAB5120-B45169. Sequences AAB5170-B45225 represent alternative polypeptides encoded by the genes, and amino acid sequences to which they are homologous. The genes and proteins have activities dependent on the tissues and cells in which they are expressed. Examples of their activities include immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytosolic; cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial; vitucide; fungicide; and ophthalmological. The secreted proteins, polynucleotides, antagonists and agonists may be useful in treating, preventing and/or diagnosing diseases and disorders such as autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. CC The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. AAC80522-C80530 and AAB45119 represent sequences used in the isolation and characterisation of the genes and proteins of the invention.			
CC				
SC	Sequence	106 AA;		
QY	Query Match	43.3%: Score 220; DB 21; Length 106;		
	Best Local Similarity	51.3%: Prod. No. 3.6e-20;		
Matches	39: Conservative	15; Mismatches	22; Indels	0; Gaps
QY	19 LFIAGVAGMCPFWLMLSEIFPLHLKGVATGVCLTNMFMAFLVKRFSIMEILRPYG 78 : : : : : : : : : : : 2 LFIMGVAYGWPGITWLIMSEVLPIRTARGVSGLVCIAASWTALVTIKSFTLPVVSTIGLY 61 : : : : : : : : : : : 79 AFULTAAFCILSVLETF 94 : : : : : : : : : : :			
Dd	62 pfffaalclysvltc 77			
RESULT	8			
ID	AAY95032 standard; Protein: 147 AA.			
XX	AAY95032:			
DT	19-JUN-2000 (first entry)			
DE	Human clone vqj_1 insertional variant ORF, SEQ ID NO:132.			
KM	Human: secreted protein; cancer: tumour; cardiovascular disorder: blood disorder: haemophilia; autoimmune disease: diabetes; inflammation: infection; fungal; bacterial; viral; HIV; allergy: arthritis; neurodegenerative disease: asthma; contraceptive: open reading frame:			

KW		ORF; variant.
XX		
OS	Homo sapiens.	
XX		
PN	WO200011015-AI.	
XX		
PD	02-MAR-2000.	
XX		
PF	24-AUG-1999;	99MO-USJ9351.
XX		
PR	24-AUG-1998;	98US--0097638.
PR	24-AUG-1998;	98US--0097659.
PR	09-SEP-1998;	98US--0099618.
PR	28-SEP-1998;	98US--0102092.
PR	25-NOV-1998;	98US--0109978.
PR	23-DEC-1998;	98US--0113645.
PR	23-DEC-1998;	98US--0113646.
PR	23-AUG-1999;	98US--0379246.
PA	(ALPH-) ALPBASE INC.	
XX		
PI	Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;	
XX		
DR	WPt: 2000-224657/19.	
XX		
PT	New secreted or transmembrane proteins and polynucleotides encoding them, useful for treating neurodegenerative disorders, autoimmune diseases and cancer -	
PS	Disclosure: Page 351-352: 357pp: English.	
XX		
CC	The invention relates to 40 human secreted proteins (AA94981-Y95020), and cDNA sequences encoding them (AAA23423-A23462). The secreted proteins of the invention include those that are thought to be only partially secreted, i.e., transmembrane proteins. The proteins of the invention may exhibit one or more activities selected from the following: cytokine activity; cell proliferation; differentiation; immune modulation; haematopoiesis regulation; tissue growth activity; activator/inhibitor activity; chemotactic/chemokinetic activity; haemostatic and thrombolytic activity; anti-inflammatory activity; and tumour inhibition activity. The proteins may be administered to patients as vaccines, and the nucleotides may be used as part of a gene therapy regime. Diseases or conditions that may be treated using the protein or nucleotide of the invention include autoimmune diseases; genetic disorders; hemophilia; cardiovascular diseases; cancer; bacterial, fungal and viral infections, especially HIV; multiple sclerosis; rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome; insulin dependent diabetes mellitus; and allergic reactions such as asthma and anaemia. They may also be used for treating wounds, burns, ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's disease, Parkinson's disease, Huntington's disease and amyotropic lateral sclerosis (ALS). Proteins with activator/inhibitor activity may additionally be useful as contraceptives. Nucleic acid sequences of the invention may be used in chromosome mapping, and as a source of diagnostic primers and probes. Sequences AA95024-Y95025, AA95029 and AA95031-Y95032 represent additional open reading frames (ORFs) that are encoded by deletion or insertional variants of the cDNA clones of the invention.	
SQ	Sequence 147 AA:	
OY	Query Match 43.3%; Score 220; DB 21: Length 147; Best Local Similarity 51.3%; Pred. No. 5.4e+20; Matches 39; Conservative 15; Mismatches 22; Indels 0; Gaps 0	
DG	19 LFIAGFAVGMPPLWLMSSEIPFLHKGAVTCVCVTNNFMFLVKFNFSIMETILPYG 78 :: :: :: :: : :: :	
DB	2 IIFmgayvgwgitvlmseqivplrtarygaagjlcylaswltafvltksflpvstqlqy 61 : : :: ::	
OY	79 AFMLTAACIIISLVLT 94 : : :: ::	
DB	62 pffiafaiclvsilvt 77	

```

RESULT 9
AAB66937
ID AAB66937 standard; Protein: 507 AA.
XX
AC AAB66937;
XX
DT 17-APR-2001 (first entry)
XX
DE Human GLUTX3.
XX
KW Human; GLUTX: gene therapy; vaccine; hexose transport modulator;
KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS Homo sapiens.
XX
PN WO200104145-A2.
XX
PD 18-JAN-2001.
XX
PE 14-JUL-2000; 2000WO-1B01042.
XX
PR 14-JUL-1999; 99US-0143907.
PR 27-AUG-1999; 99US-0151140.
PR 23-FEB-2000; 2000US-0184285.
PR 13-JUL-2000; 2000US-0616132.
XX
PA (UYLA-) UNIV LAUSANNE.
XX
PI Thorens B, Ibberson M, Uldry M;
XX
DR WPI: 2001-112615/12.
DR N-PSDB; AAF55870.
XX
PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
PT e.g. ischaemia and diabetes -
XX
PS Claim 11; Page 81-82; 124pp; English.
XX
CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischaemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is human GLUTX3.
XX
SQ Sequence 507 AA;

Query Match 42.7%; Score 217; DB 22; Length 507;
Best Local Similarity 51.3%; Pred. No. 5.7e-19;
Matches 39; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

OY 19 LFIAGFAVGMPIPWLMSEIFPLHIKGVATGVCVLTNMFMAFLVTKFNSIMEILRPYG 78
DB 404 lfimgyavgwptlwlmselvlplargvasglcvlaslatafvltskflpvsctfglqv 463
OY 79 AFWLTAFCILSVLFT 94
DB 464 pflffaalcivslvft 479

RESULT 10
AAB66941
ID AAB66941 standard; Protein: 507 AA.
XX
AC AAB66941;
XX
DT 17-APR-2001 (first entry)

```

```

XX
DE GLUTX3 consensus sequence.
XX
KW GLUTX: gene therapy; vaccine; hexose transport modulator; human; rat;
KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS Homo sapiens.
XX
PN WO200104145-A2.
XX
PD 18-JAN-2001.
XX
PE 14-JUL-2000; 2000WO-1B01042.
XX
PR 14-JUL-1999; 99US-0143907.
PR 27-AUG-1999; 99US-0151140.
PR 23-FEB-2000; 2000US-0184285.
PR 13-JUL-2000; 2000US-0616132.
XX
PA (UYLA-) UNIV LAUSANNE.
XX
PI Thorens B, Ibberson M, Uldry M;
XX
DR WPI: 2001-112615/12.
DR N-PSDB; AAF55870.
XX
PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
PT e.g. ischaemia and diabetes -
XX
PS Claim 11; Page 83-84; 124pp; English.
XX
CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischaemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is a consensus sequence
CC for GLUTX3.
XX
SQ Sequence 507 AA;

Query Match 42.7%; Score 217; DB 22; Length 507;
Best Local Similarity 51.3%; Pred. No. 5.7e-19;
Matches 39; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

OY 19 LFIAGFAVGMPIPWLMSEIFPLHIKGVATGVCVLTNMFMAFLVTKFNSIMEILRPYG 78
DB 404 lfimgyavgwptlwlmselvlplargvasglcvlaslatafvltskflpvsctfglqv 463
OY 79 AFWLTAFCILSVLFT 94
DB 464 pflffaalcivslvft 479

RESULT 11
AAB66938
ID AAB66938 standard; Protein: 503 AA.
XX
AC AAB66938;
XX
DT 17-APR-2001 (first entry)
XX
DE Rat GLUTX3.
XX
KW Rat; GLUTX: gene therapy; vaccine; hexose transport modulator;
KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS Rattus sp.

```

XX WO200104145-A2.
 XX
 PN 18-JAN-2001.
 XX
 PD
 PF 14-JUL-2000; 2000WO-1B01042.
 XX
 XX 14-JUL-1999; 99US-0143907.
 PR 27-AUG-1999; 99US-0151140.
 PR 23-FEB-2000; 2000US-0184285.
 PR 13-JUL-2000; 2000US-0616132.
 XX
 XX (UYLA-) UNIV LAUSANNE.
 PA
 XX Thorens B, Ibberson M, Uldry M.
 PI
 DR MPI: 2001-112615/12.
 DR N-PSDB; AAF55871.
 XX
 XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -
 XX
 PS Claim 11: Page 82-83; 124pp: English.
 XX
 XX The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB6693-AAF66941). The GLUTX proteins are related to the facilitative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is rat GLUTX3.
 XX
 SQ Sequence 503 AA:
 QY 2 PAVVHGLAMLAAGSMCLFAGFVAGWGPIPMILMSEIFPHIKGVATGCVLTNMFMAF 61
 Db 383 paaafytlilpllatlunltlmgyamgwpitwllmsevlprargvasgiclvswlatf 442
 QY 62 LVTKFENSIMEILRPYGAEFWLTAFCILSVLEFT 94
 Db 443 vlkylflavnaftglqvffiffsaicllslft 475

RESULT 12
 AAM25722
 ID AAM25722 standard; Protein: 70 AA.
 XX
 AC AAM25722:
 XX
 DT 16-OCT-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:1237.
 XX
 KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
 KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
 KW antibacterial; endocrine; cardiant; central nervous system; virucide;
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antinaemic; anaemia;
 KW antiaggregant; haemostatic; vulnery; antitumor; osteopathic; eczema;
 KW dermatological; antiallergic; antisthmatic; antidiabetic; cytostatic;
 KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;

KW neurological disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200153455-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 22-DEC-2000; 2000WO-US35017.
 XX
 XX 23-DEC-1999; 99US-0471275.
 PR 21-JAN-2000; 2000US-048725.
 PR 25-APR-2000; 2000US-0552317.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Drmanac RT;
 PI
 DR MPI: 2001-457603/49.
 DR N-PSDB; AAH99663.
 XX
 XX Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 PT
 PS Claim 20: Page 256; 1217pp: English.
 XX
 XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: antiinflammatory; antirheumatic;
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antinaemic; antiaggregant; haemostatic; vulnery;
 CC antitumor; osteopathic; dermatological; antiallergic; antisthmatic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.
 CC
 XX
 SQ Sequence 70 AA:
 QY 12 LAVGSMCLFAGFVAGWGPIPMILMSEIFPHIKGVATGCVLTNMFMAFVTKFEF 67
 Db 1 lqgrkvllfsvgyvvgwptwllmsevlprargvasgiclvswlatfsltsf 56

RESULT 13
 AAG37400
 ID AAG37400 standard; Protein: 193 AA.
 XX
 AC AAG37400:
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 45980.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132866.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140053.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 99US-0140823.

PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149176.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 99US-0153070.

PR	13-SEP-1999;	990S-0153758.
PR	13-SEP-1999;	990S-0154018.
PR	16-SEP-1999;	990S-0154039.
PR	20-SEP-1999;	990S-0154779.
PR	22-SEP-1999;	990S-0155139.
PR	23-SEP-1999;	990S-0155486.
PR	24-SEP-1999;	990S-0156589.
PR	28-SEP-1999;	990S-0156458.
PR	29-SEP-1999;	990S-0156596.
PR	04-OCT-1999;	990S-0157117.
PR	05-OCT-1999;	990S-0157753.
PR	06-OCT-1999;	990S-0157865.
PR	07-OCT-1999;	990S-0158029.
PR	08-OCT-1999;	990S-0158232.
PR	12-OCT-1999;	990S-0158369.
PR	13-OCT-1999;	990S-0159293.
PR	13-OCT-1999;	990S-0159294.
PR	13-OCT-1999;	990S-0159295.
PR	14-OCT-1999;	990S-0159329.
PR	14-OCT-1999;	990S-0159330.
PR	14-OCT-1999;	990S-0159331.
PR	14-OCT-1999;	990S-0159637.
PR	14-OCT-1999;	990S-0159584.
PR	18-OCT-1999;	990S-0159588.
PR	21-OCT-1999;	990S-0160741.
PR	21-OCT-1999;	990S-0160767.
PR	21-OCT-1999;	990S-0160768.
PR	21-OCT-1999;	990S-0160770.
PR	21-OCT-1999;	990S-0160814.
PR	21-OCT-1999;	990S-0160815.
PR	22-OCT-1999;	990S-0160980.
PR	22-OCT-1999;	990S-0160981.
PR	22-OCT-1999;	990S-0160989.
PR	25-OCT-1999;	990S-0161404.
PR	25-OCT-1999;	990S-0161405.
PR	25-OCT-1999;	990S-0161406.
PR	26-OCT-1999;	990S-0161359.
PR	26-OCT-1999;	990S-0161360.
PR	26-OCT-1999;	990S-0161361.
PR	28-OCT-1999;	990S-0161920.
PR	28-OCT-1999;	990S-0161992.
PR	28-OCT-1999;	990S-0161993.
PR	29-OCT-1999;	990S-0162142.
Query Match 36.9%; Score 187.5; DB 21; Length 193;		
Best Local Similarity 43.2%; Pred. No. 9.2e-16;		
Matches 38; Conservative 18; Mismatches 31; Indels 1; Gaps 1;	OY	6 HGLMLAVGSGCLTACAGVGCMPILPWLMSSEIPPLHKLKGVATCVLTNNFMFLVTR 65
I : : : : : : : : : : :		85 heavyplavvglmlylgsfsagmgampwvymseifrpinkkyagmalivnwfgawasy 144
Db	OY	66 EFNISMEILRPYGAFWLTAAFCILSLVF 93
		I : I : : : :
DB	145 tnfIms-wssygtfIIyaalnaIaIf 171	
RESULT 14		
AAG37422	ID	AAG37422 standard; Protein: 193 AA.
XX AC AAG37422:		
XX DT 18-OCT-2000 (first entry)		
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 46010.		
XX KW Protein identification; signal transduction pathway; metabolic pathway;		
KM hybridisation assay; genetic mapping; gene expression control; promoter;		
KX termination sequence.		
EX Arabidopsis thaliana.		
XX		

PM	EPI01033405-A2.	
XX	06-SEP-2000.	
XX		
PF	25-FEB-2000.	2000EP-0301439
XX		
PR	25-FEB-1999.	9905-0121825
PR	05-MAR-1999.	9905-0123180
PR	23-MAR-1999.	9905-0123548
PR	03-MAR-1999.	9905-0125758
PR	25-MAR-1999.	9905-0126264
PR	29-MAR-1999.	9905-0126785
PR	01-APR-1999.	9905-0127462
PR	06-APR-1999.	9905-0128234
PR	08-APR-1999.	9905-0128714
PR	16-APR-1999.	9905-0129845
PR	19-APR-1999.	9905-0130077
PR	21-APR-1999.	9905-0130449
PR	23-APR-1999.	9905-0130510
PR	28-APR-1999.	9905-0130891
PR	30-APR-1999.	9905-0131448
PR	04-MAY-1999.	9905-0132407
PR	05-MAY-1999.	9905-0132484
PR	06-MAY-1999.	9905-0132485
PR	06-MAY-1999.	9905-0132486
PR	07-MAY-1999.	9905-0132487
PR	11-MAY-1999.	9905-0133863
PR	14-MAY-1999.	9905-0134218
PR	14-MAY-1999.	9905-0134219
PR	14-MAY-1999.	9905-0134221
PR	18-MAY-1999.	9905-0134370
PR	18-MAY-1999.	9905-0134376
PR	19-MAY-1999.	9905-0134941
PR	20-MAY-1999.	9905-0135314
PR	21-MAY-1999.	9905-0135323
PR	24-MAY-1999.	9905-0135629
PR	25-MAY-1999.	9905-0136021
PR	27-MAY-1999.	9905-0136392
PR	28-MAY-1999.	9905-0136782
PR	01-JUN-1999.	9905-0137222
PR	03-JUN-1999.	9905-0137528
PR	04-JUN-1999.	9905-0137502
PR	07-JUN-1999.	9905-0137724
PR	08-JUN-1999.	9905-0138094
PR	10-JUN-1999.	9905-0138540
PR	14-JUN-1999.	9905-0138847
PR	16-JUN-1999.	9905-0139119
PR	16-JUN-1999.	9905-0139453
PR	16-JUN-1999.	9905-0139455
PR	17-JUN-1999.	9905-0139492
PR	18-JUN-1999.	9905-0139454
PR	18-JUN-1999.	9905-0139455
PR	18-JUN-1999.	9905-0139456
PR	18-JUN-1999.	9905-0139457
PR	18-JUN-1999.	9905-0139458
PR	18-JUN-1999.	9905-0139459
PR	18-JUN-1999.	9905-0139460
PR	18-JUN-1999.	9905-0139461
PR	18-JUN-1999.	9905-0139462
PR	18-JUN-1999.	9905-0139463
PR	18-JUN-1999.	9905-0139464
PR	18-JUN-1999.	9905-0139465
PR	18-JUN-1999.	9905-0139466
PR	18-JUN-1999.	9905-0139467
PR	18-JUN-1999.	9905-0139468
PR	18-JUN-1999.	9905-0139469
PR	18-JUN-1999.	9905-0139470
PR	18-JUN-1999.	9905-0139471
PR	18-JUN-1999.	9905-0139472
PR	18-JUN-1999.	9905-0139473
PR	18-JUN-1999.	9905-0139474
PR	18-JUN-1999.	9905-0139475
PR	18-JUN-1999.	9905-0139476
PR	18-JUN-1999.	9905-0139477
PR	18-JUN-1999.	9905-0139478
PR	18-JUN-1999.	9905-0139479
PR	18-JUN-1999.	9905-0139480
PR	18-JUN-1999.	9905-0139481
PR	18-JUN-1999.	9905-0139482
PR	18-JUN-1999.	9905-0139483
PR	18-JUN-1999.	9905-0139484
PR	18-JUN-1999.	9905-0139485
PR	18-JUN-1999.	9905-0139486
PR	18-JUN-1999.	9905-0139487
PR	18-JUN-1999.	9905-0139488
PR	18-JUN-1999.	9905-0139489
PR	18-JUN-1999.	9905-0139490
PR	18-JUN-1999.	9905-0139491
PR	18-JUN-1999.	9905-0139492
PR	18-JUN-1999.	9905-0139493
PR	18-JUN-1999.	9905-0139494
PR	18-JUN-1999.	9905-0139495
PR	18-JUN-1999.	9905-0139496
PR	18-JUN-1999.	9905-0139497
PR	18-JUN-1999.	9905-0139498
PR	18-JUN-1999.	9905-0139499
PR	18-JUN-1999.	9905-0139500
PR	18-JUN-1999.	9905-0139501
PR	18-JUN-1999.	9905-0139502
PR	18-JUN-1999.	9905-0139503
PR	18-JUN-1999.	9905-0139504
PR	18-JUN-1999.	9905-0139505
PR	18-J	

```

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

```

```

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 36.3%; Score 187.5; DB 21; Length 193;
 Best Local Similarity 43.2%; Pred. No. 9.2e-16;
 Matches 38; Conservative 18; Mismatches 31; Indels 1; Gaps 1;

```

QY 6 HLGLAWLAVGSMCLFIFAGFVAVGCPFIPWLMSEIFPHIGVATGVCVLTNNMFAFLVTK 65
D 85 heavytlaavvglmvyigstfsagmpwvmselfipnikyvgmatlvnwfagawvsv 144
QY 66 EFNSIMEILRPYGAFWLTAAFCILSVLF 93
D 145 tftflms-wssygtfllyaainalatayf 171

```

```

RESULT 15
AAG37398
ID AAG37398 standard; Protein; 212 AA.
XX
AC AAG37398;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 45978.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.

```

XX 25-FEB-2000: 2000EP-0301439.
PR 25-FEB-1999: 99US-0121825.
XX 05-MAR-1999: 99US-0123180.
PR 09-MAR-1999: 99US-0123548.
PR 23-MAR-1999: 99US-0125788.
PR 25-MAR-1999: 99US-0126264.
PR 29-MAR-1999: 99US-0126785.
PR 01-APR-1999: 99US-0127462.
PR 06-APR-1999: 99US-0128234.
PR 08-APR-1999: 99US-0128714.
PR 16-APR-1999: 99US-0129845.
PR 19-APR-1999: 99US-0130077.
PR 21-APR-1999: 99US-0130449.
PR 23-APR-1999: 99US-0130510.
PR 28-APR-1999: 99US-0130891.
PR 30-APR-1999: 99US-0131449.
PR 30-APR-1999: 99US-0132048.
PR 04-MAY-1999: 99US-0132407.
PR 05-MAY-1999: 99US-0132484.
PR 06-MAY-1999: 99US-0132485.
PR 06-MAY-1999: 99US-0132486.
PR 07-MAY-1999: 99US-0132487.
PR 11-MAY-1999: 99US-0134256.
PR 14-MAY-1999: 99US-0134218.
PR 14-MAY-1999: 99US-0134219.
PR 14-MAY-1999: 99US-0134221.
PR 14-MAY-1999: 99US-0134370.
PR 18-MAY-1999: 99US-0134768.
PR 19-MAY-1999: 99US-0134941.
PR 20-MAY-1999: 99US-0135124.
PR 21-MAY-1999: 99US-0135353.
PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136592.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137528.
PR 04-JUN-1999: 99US-0137528.
PR 07-JUN-1999: 99US-0137724.
PR 08-JUN-1999: 99US-0138094.
PR 10-JUN-1999: 99US-0138540.
PR 10-JUN-1999: 99US-0138847.
PR 14-JUN-1999: 99US-0139119.
PR 16-JUN-1999: 99US-0139452.
PR 16-JUN-1999: 99US-0139453.
PR 17-JUN-1999: 99US-0139492.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139763.
PR 21-JUN-1999: 99US-0139817.
PR 22-JUN-1999: 99US-0139819.
PR 23-JUN-1999: 99US-0140353.
PR 23-JUN-1999: 99US-0140354.
PR 24-JUN-1999: 99US-0140695.
PR 28-JUN-1999: 99US-0140823.
PR 29-JUN-1999: 99US-0140891.
PR 30-JUN-1999: 99US-0141287.
PR 01-JUL-1999: 99US-0141842.
PR 01-JUL-1999: 99US-0142154.
PR 02-JUL-1999: 99US-0142055.
PR 06-JUL-1999: 99US-0142390.

PR 08-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142920.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143542.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144632.
PR 20-JUL-1999: 99US-0144884.
PR 21-JUL-1999: 99US-0144814.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
PR 22-JUL-1999: 99US-0145192.
PR 23-JUL-1999: 99US-0145145.
PR 23-JUL-1999: 99US-0145218.
PR 23-JUL-1999: 99US-0145224.
PR 26-JUL-1999: 99US-0145276.
PR 27-JUL-1999: 99US-0145913.
PR 27-JUL-1999: 99US-0145918.
PR 27-JUL-1999: 99US-0145919.
PR 28-JUL-1999: 99US-0145951.
PR 02-AUG-1999: 99US-0146386.
PR 02-AUG-1999: 99US-0146388.
PR 03-AUG-1999: 99US-0146389.
PR 03-AUG-1999: 99US-0147028.
PR 04-AUG-1999: 99US-0147204.
PR 04-AUG-1999: 99US-0147302.
PR 05-AUG-1999: 99US-0147312.
PR 05-AUG-1999: 99US-0147260.
PR 06-AUG-1999: 99US-0147303.
PR 06-AUG-1999: 99US-0147416.
PR 09-AUG-1999: 99US-0147493.
PR 09-AUG-1999: 99US-0147935.
PR 10-AUG-1999: 99US-0148171.
PR 11-AUG-1999: 99US-0148319.
PR 12-AUG-1999: 99US-0148341.
PR 13-AUG-1999: 99US-0148565.
PR 13-AUG-1999: 99US-0148684.
PR 16-AUG-1999: 99US-0149368.
PR 17-AUG-1999: 99US-0149175.
PR 18-AUG-1999: 99US-0149426.
PR 20-AUG-1999: 99US-0149722.
PR 20-AUG-1999: 99US-0149723.
PR 20-AUG-1999: 99US-0149729.
PR 23-AUG-1999: 99US-0149902.
PR 23-AUG-1999: 99US-0149930.
PR 25-AUG-1999: 99US-0150566.
PR 26-AUG-1999: 99US-0150884.
PR 27-AUG-1999: 99US-0151065.
PR 27-AUG-1999: 99US-0151066.
PR 27-AUG-1999: 99US-0151080.
PR 30-AUG-1999: 99US-0151303.
PR 31-AUG-1999: 99US-0151438.
PR 01-SEP-1999: 99US-0151930.
PR 07-SEP-1999: 99US-0152363.
PR 10-SEP-1999: 99US-0153070.
PR 13-SEP-1999: 99US-0153758.
PR 15-SEP-1999: 99US-0154018.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.

PR	24-SEP-1999;	9905-0155659.
PR	28-SEP-1999;	9905-0156458.
PR	29-SEP-1999;	9905-0156596.
PR	04-OCT-1999;	9905-0157117.
PR	05-OCT-1999;	9905-0157753.
PR	06-OCT-1999;	9905-0157865.
PR	07-OCT-1999;	9905-0158029.
PR	08-OCT-1999;	9905-0158332.
PR	12-OCT-1999;	9905-0158369.
PR	13-OCT-1999;	9905-0159293.
PR	13-OCT-1999;	9905-0159294.
PR	13-OCT-1999;	9905-0159295.
PR	14-OCT-1999;	9905-0159329.
PR	14-OCT-1999;	9905-0159330.
PR	14-OCT-1999;	9905-0159331.
PR	14-OCT-1999;	9905-0159637.
PR	14-OCT-1999;	9905-0159638.
PR	18-OCT-1999;	9905-0159584.
PR	21-OCT-1999;	9905-0160741.
PR	21-OCT-1999;	9905-0160767.
PR	21-OCT-1999;	9905-0160768.
PR	21-OCT-1999;	9905-0160770.
PR	21-OCT-1999;	9905-0160814.
PR	21-OCT-1999;	9905-0160815.
PR	22-OCT-1999;	9905-0160980.
PR	22-OCT-1999;	9905-0160981.
PR	22-OCT-1999;	9905-0160989.
PR	25-OCT-1999;	9905-0161404.
PR	25-OCT-1999;	9905-0161405.
PR	25-OCT-1999;	9905-0161406.
PR	26-OCT-1999;	9905-0161359.
PR	26-OCT-1999;	9905-0161360.
PR	26-OCT-1999;	9905-0161361.
PR	28-OCT-1999;	9905-0161920.
PR	28-OCT-1999;	9905-0161992.
PR	28-OCT-1999;	9905-0161993.
PR	29-OCT-1999;	9905-0162142.

Query Match	36.98;	Score 187.5;	DB 21;	Length 212;
Best Local Similarity	43.38;	Prod No 18-15;		

Matches 38; Conservative 18; Mismatches 31; Indels 1; Caps 1;

6 HGLAWLVGSMCLFIAGFAVGMQPIPWLLMSEIFPLHIKGVATGVCLTNEMAEFLVTK 65

```

104  heapv1avdqimvviqsfasmqamowvymseifrijkavraqamatjvnvfravavsv 163
      | : ||| : :: | : ||::|||:::||| | : | ||| | : |

```

66 FENSI METIRPYGAEMITAECTISWE Q3

[illegible]

104 LINDHUS WSSYGRILLYAANALAIVI 190

Search completed: February 13, 2002, 21:50:12
Job time: 6263 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 21:51:37 ; Search time 72.65 Seconds
(without alignments)
29.116 Million cell updates/sec

Title: US-09-516-493-12

Perfect score: 508

Sequence: 1 EPADVHLGLAWLAVGSMCLF.....RPYGAFWLTAACILSVLFT 94

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCRTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	130.5	25.7	493	2	US-09-031-392-10
2	130.5	25.7	493	4	US-09-299-549-10
3	130	25.6	524	2	US-08-928-692-12
4	129.5	25.4	484	2	US-08-928-692-13
5	124	24.4	488	2	US-08-928-692-10
6	123	24.2	494	4	US-09-031-392-5
7	123	24.2	494	4	US-09-299-549-5
8	122.5	24.1	534	2	US-09-031-392-4
9	122.5	24.1	534	4	US-09-299-549-4
10	122	24.0	583	4	US-09-031-392-3
11	122	24.0	583	4	US-09-299-549-3
12	120.5	23.7	509	2	US-09-031-392-6
13	120.5	23.7	509	4	US-09-299-549-6
14	113	22.2	492	2	US-08-355-844-3
15	113	22.2	492	5	PCT-US95-16126-3
16	95	18.7	488	2	US-08-928-692-11
17	84.5	16.6	563	2	US-09-031-392-2
18	84.5	16.6	563	4	US-09-299-549-2
19	80.5	15.8	500	4	US-09-031-392-7
20	80.5	15.8	500	4	US-09-299-549-7
21	78.5	15.5	532	3	US-08-948-564-10
22	68.5	13.5	422	3	US-08-492-459-2
23	68.5	13.5	422	3	US-08-492-459-4
24	68.5	13.5	422	3	US-08-423-752-2
25	68.5	13.5	422	3	US-08-423-752-4
26	68.5	13.5	422	4	US-08-716-873-7
27	68.5	13.5	422	4	US-08-716-873-16

28	68.5	13.5	422	4	US-08-716-873-18	Sequence 18, Appl
29	68.5	13.5	422	4	US-09-368-431-7	Sequence 7, Appl
30	68.5	13.5	422	4	US-09-368-431-16	Sequence 16, Appl
31	68.5	13.5	422	4	US-09-368-431-18	Sequence 18, Appl
32	68.5	13.5	430	1	US-07-937-609-23	Sequence 23, Appl
33	68.5	13.5	430	2	US-08-919-624-3	Sequence 3, Appl
34	68.5	13.5	430	4	US-08-029-170-23	Sequence 23, Appl
35	68.5	13.5	450	1	US-07-937-609-24	Sequence 24, Appl
36	68.5	13.5	488	4	US-08-029-170-24	Sequence 24, Appl
37	68.5	13.5	488	4	US-08-776-511-2	Sequence 2, Appl
38	68.5	13.5	508	2	US-08-850-293-5	Sequence 5, Appl
39	67.5	13.3	428	1	US-08-570-157-5	Sequence 5, Appl
40	67.5	13.3	428	4	US-08-029-170-31	Sequence 31, Appl
41	67.5	13.3	444	1	US-07-937-609-14	Sequence 14, Appl
42	67.5	13.3	444	4	US-08-029-170-14	Sequence 14, Appl
43	64.5	12.7	401	3	US-08-492-459-8	Sequence 8, Appl
44	64.5	12.7	401	3	US-08-492-459-28	Sequence 8, Appl
45	64.5	12.7	401	3	US-08-423-752-8	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-09-031-392-10
; Sequence 10, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Melkielejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 493 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-031-392-10

Query Match 25.7%; Score 130.5; DB 2; Length 493;
Best Local Similarity 33.7%; Pred. No. 2.7e-08;
Matches 29; Conservative 20; Mismatches 36; Indels 1; Gaps 1;
Oy 9 LANLAVGSMCLFAGFVWGCPIPMILMSLFIPIHIGVATGCVLTNNMFMAFLVTKFEN 68
Db 365 MSYVSIVAIRGFVAFVFGCPIDPMFIVALELFSQGRPRPAIIVAGFSNMTSNFTVGLLFQ 424


```

1  COUNTRY:  USA
2  ZIP:  02110-2804
3  COMPUTER READABLE FORM:
4  MEDIUM TYPE:  Diskette
5  COMPUTER:  IBM Compatible
6  OPERATING SYSTEM:  Windows95
7  SOFTWARE:  FastSeq for Windows Version 2.0
8  CURRENT APPLICATION DATA:
9  APPLICATION NUMBER:  US/09/031.392
10 FILING DATE:  26-FEB-1998
11 ATTORNEY/AGENT INFORMATION:
12 NAME:  Melkielehn, Ph.D., Anita L.
13 REGISTRATION NUMBER:  35,283
14 REFERENCE/DOCKET NUMBER:  07334/072001
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE:  617/542-5070
17 TELEFAX:  617/542-8906
18 TELEX:  200154
19 INFORMATION FOR SEQ ID NO:  4:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH:  534 amino acids
22 TYPE:  amino acid
23 TOPOLOGY:  linear
24 MOLECULE TYPE:  protein
25
26 US-09-031-392-4

```

[illegible]

TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,392
FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Melkielejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 509 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-031-392-6

Query Match 23.7%; Score 120.5; DB 2; Length 509;
Best Local Similarity 29.1%; Pred. No. 4.9e-07;
Matches 25; Conservative 22; Mismatches 38; Indels 1; Gaps 1;

QY 9 LAWLAVGSMCLFIAGFVAGWGPILPWLMSLIFPLHKGATGCVLTNWMFAFLVTKFN 68
DB 380 MSYVSIVAIIFGVAFEEIGPDPIMFIVAELEFSOGPRPAAVAAGFSNWTSTNFIIGMGFQ 439
QY 69 SIMELIRPYGAFMLTAFCILSVLFT 94
DB 440 YVAEAMGPY-VLLFAVLLGFIIFT 464

RESULT 13
US-09-299-549-6
Sequence 6, Application US/09299549
Patent No. 6136547
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
APPLICANT: Meng, Xun
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
TITLE OF INVENTION: GLUTEX AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,549
FILING DATE: 26-APR-1999
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/031,392
FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Melkielejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 509 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-299-549-6

Query Match 23.7%; Score 120.5; DB 4; Length 509;
Best Local Similarity 29.1%; Pred. No. 4.9e-07;
Matches 25; Conservative 22; Mismatches 38; Indels 1; Gaps 1;

QY 9 LAWLAVGSMCLFIAGFVAGWGPILPWLMSLIFPLHKGATGCVLTNWMFAFLVTKFN 68
DB 380 MSYVSIVAIIFGVAFEEIGPDPIMFIVAELEFSOGPRPAAVAAGFSNWTSTNFIIGMGFQ 439
QY 69 SIMELIRPYGAFMLTAFCILSVLFT 94
DB 440 YVAEAMGPY-VLLFAVLLGFIIFT 464

RESULT 14
US-08-355-844-3
Sequence 3, Application US/08355844
Patent No. 5940307
GENERAL INFORMATION:
APPLICANT: Fischbarg, Jorge
APPLICANT: Czegledy, Ferenc
APPLICANT: Iserovich, Pavel
APPLICANT: Li, Jun
APPLICANT: Cheung, Min
TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN
TITLE OF INVENTION: STRUCTURE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/355,844
FILING DATE: 14-DEC-1994
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Tang, Henry Y.S.
REGISTRATION NUMBER: 29,705
REFERENCE/DOCKET NUMBER: A29927-50/29910
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2586
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 492 amino acids
TYPE: amino acid
STRANDEDNESS: single

```

;
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Human
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..492
; OTHER INFORMATION: Facilitative glucose transportor
; OTHER INFORMATION: Glut1 protein
US-08-355-844-3

Query Match          22.2%; Score 113; DB 2; Length 492;
Best local Similarity 30.6%; Pred. No. 4.1e-06;
Matches 26; Conservative 20; Mismatches 35; Indels 4; Gaps 1;

OY 9 LAMLVGSMCLFIAGFAVGMGPIPMILMSEIFPLHIKGVATGCVLTNNFMALVTKERN 68
    :::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 364 MSYLSIVAIKGFVAFVFGPGPIPMFIVAELESQGRPAIAVAGFSNMTSNFTVGMCFQ 423
    :::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY 69 SIMELIRPYGAFWLTAAFCILSYLF 93
    :::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 424 YVEQLCGPY---VFILFTVLVLF 444

RESULT 15
PCT-US95-16126-3
; Sequence 3, Application PC/TUS9516126
; GENERAL INFORMATION:
; APPLICANT: Fischbarg, Jorge
; APPLICANT: Czegledy, Ferenc
; APPLICANT: Iserovich, Pavel
; APPLICANT: Li, Jun
; APPLICANT: Cheung, Min
; TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16126
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/355,844
; FILING DATE: 14-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Tang, Henry Y. S.
; REGISTRATION NUMBER: 29,705
; REFERENCE/DOCKET NUMBER: A29927-50/29910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-765-2519
; TELEFAX: 212-765-2586
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 492 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Human
; FEATURE:
```

```

;
; NAME/KEY: Peptide
; LOCATION: 1..492
; OTHER INFORMATION: Facilitative glucose transportor
; OTHER INFORMATION: Glut1 protein
PCT-US95-16126-3

Query Match          22.2%; Score 113; DB 5; Length 492;
Best local Similarity 30.6%; Pred. No. 4.1e-06;
Matches 26; Conservative 20; Mismatches 35; Indels 4; Gaps 1;

OY 9 LAMLVGSMCLFIAGFAVGMGPIPMILMSEIFPLHIKGVATGCVLTNNFMALVTKERN 68
    :::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 364 MSYLSIVAIKGFVAFVFGPGPIPMFIVAELESQGRPAIAVAGFSNMTSNFTVGMCFQ 423
    :::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY 69 SIMELIRPYGAFWLTAAFCILSYLF 93
    :::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 424 YVEQLCGPY---VFILFTVLVLF 444
```

Search completed: February 13, 2002, 21:51:38
Job time: 6265 sec

Thu Feb 14 07:44:20 2002

us-09-516-493-12.rail

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 21:53:11 ; Search time 79.58 Seconds

(without alignments)
89.977 Million cell updates/sec

Title: US-09-516-493-12

Perfect score: 508

Sequence: 1 EPADVHLGLAVLAVGSMCLF.....RPYGAFWLTAFCILSVLFT 94

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	187.5	36.9	348	2 D84922	probable sugar tra
2	171.5	33.8	487	2 E96782	hypothetical prote
3	162.5	32.0	516	2 T12199	monosaccharid tran
4	159.5	31.4	514	2 S25009	monosaccharide tra
5	159.5	31.4	523	2 S25015	monosaccharide tra
6	158.5	31.2	376	2 T43176	probable myo-inosi
7	158	31.1	493	2 F96696	protein FIN21.12 (
8	158	31.1	521	2 G84864	probable membrane
9	157.5	31.0	522	2 S12042	glucose transport
10	157.5	31.0	522	2 E86246	glucose transport
11	157	30.9	467	2 B96829	probable sugar tra
12	155.5	30.6	490	2 T14545	hexose transport
13	155.5	30.6	534	2 S38435	hexose transport p
14	155	30.5	457	2 E70070	metabolite transpo
15	154.5	30.4	483	2 D96589	hypothetical prote
16	154.5	30.4	523	2 T10122	hexose transport p
17	152.5	30.0	534	2 S14144	hexose transport p
18	151.5	29.8	433	2 G86812	D-xylose proton-sy
19	151	29.7	582	2 F71431	hypothetical prote
20	150.5	29.6	580	2 D86426	hypothetical prote
21	150	29.5	734	2 H86340	Sugar transporter
22	149.5	29.4	557	2 T81235	myo-inositol trans
23	146.5	28.8	472	2 S47089	arabinose-protom s
24	146	28.7	472	2 B26430	L-arabinose isomer
25	146	28.7	472	2 E85936	hypothetical prote
26	146	28.7	508	2 T05156	probable glucose t
27	145	28.5	613	2 T27077	hypothetical prote
28	142.5	28.1	461	2 G85059	probable sugar tra
29	142.5	28.1	729	2 T06127	probable sugar tra

30	141.5	27.9	510	2 T10124	hexose transport p
31	140.5	27.7	464	2 F65079	galactose-protom s
32	140.5	27.7	491	2 A26430	xylose transport p
33	140.5	27.7	491	2 B86096	xylose-protom symp
34	139	27.4	523	2 S06920	glucose transport
35	139	27.4	580	2 D84772	probable sugar tra
36	138.5	27.3	82	2 T10145	sugar transport pr
37	138.5	27.3	526	2 T01853	probable hexose tr
38	137.5	27.1	575	2 T43400	myo-inositol trans
39	136.5	26.9	464	2 F85951	galactose-protom s
40	136	26.8	517	2 D96539	hypothetical prote
41	135	26.6	522	2 A31556	glucose transport
42	134.5	26.5	547	2 A48442	membrane transport
43	134	26.4	522	2 T10150	sugar transport pr
44	134	26.4	606	2 T27072	hypothetical prote
45	133.5	26.3	378	2 G86220	hypothetical prote

ALIGNMENTS

RESULT 1
D84922
Probable sugar transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: D84922
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Unayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eilen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: D84922
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-348 <STO>
A:Cross-references: GB:AE002093; NID:94249409; PIDN:AMD13706.1; GSPDB:GN00139
C:Genetics:
A:Gene: AT2g48020
A:Map position: 2

Query Match 36.9%; Score 187.5; DB 2; Length 348;
Best Local Similarity 43.2%; Pred. No. 3.8e-12;
Matches 38; Conservative 18; Mismatches 31; Indels 1; Gaps 1;

OY 6 HGLAVLAVGSMCLFVAGFVAGVGPVPLVLMSEIFPLIKVATGCVLTVMFMAFLVTK 65
I : I I : : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 240 HEAVPVLAIVGIMVIGSFSAGMGAMPVVMSEIFPIKVGAGCATLVVMFGMAVSY 299

OY 66 EFNSIMELIRYGAFWLTAFCILSVLF 93
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 300 TEFNPLMS-WSSYGFLLYAAINALAIYF 326

RESULT 2
E96782
hypothetical protein F22H5.6 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: E96782
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizaf, L.
Nature 408, 816-820, 2000
C:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitli, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Schwart, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallo
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719

A:Accession: E96782
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-487 <STO>
 A:Cross-references: GB:AE005173; NID:g10092276; PIDN:AA612669.1; GSPDB:GND0141
 C:Genetics:
 A:Gene: F22H5.6
 A:Map position: 1
 C:Superfamily: glucose transport protein

Query Match 33.8%; Score 171.5; DB 2; Length 487;
 Best Local Similarity 37.4%; Pred. No. 2.4e-10;
 Matches 34; Conservative 21; Mismatches 35; Indels 1; Gaps 1;

OY 3 ADVHLGLAWLAVGSMCLFIAGFAVGMPIPWLMSEIFPLHIKGVATGCVLTNNMFMAFL 62
 DB 377 SDMTSLVLSVGVAMVVFSLGMPPLWLMSEIFPLVNIKGLAGSIATLANWFPSWL 436

OY 63 VTKEFNSIMEILRPYGAFWLTAAFCILSVLF 93
 DB 437 ITMTANLL-AMSSGFTLYGLVCATTVF 466

RESULT 3
 T12199
 monosaccharid transport protein - fava bean
 C:Species: Vicia faba (fava bean)
 C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 20-Jun-2000
 C:Accession: T12199
 R:Weber, H.; Borstjuk, L.; Helm, U.; Sauer, N.; Mobus, U.
 Plant Cell 9, 895-908, 1997
 A:Title: A role for sugar transporters during seed development: molecular characterization
 A:Reference number: Z17451; MUID:97355984
 A:Accession: T12199
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-516 <WEB>
 A:Cross-references: EMBL:293775
 C:Genetics:
 A:Gene: hexT
 C:Superfamily: glucose transport protein

Query Match 32.0%; Score 162.5; DB 2; Length 516;
 Best Local Similarity 36.6%; Pred. No. 2.2e-09;
 Matches 34; Conservative 20; Mismatches 38; Indels 1; Gaps 1;

OY 1 EPADVHLGLAWLAVGSMCLFIAGFAVGMPIPWLMSEIFPLHIKGVATGCVLTNNMFMA 60
 DB 378 EPDDLPRKMYAIVVLFICIVVAGFAWSGPLGWLVPSEIFLEIRSAQSIVNVSMLEPT 437
 OY 61 PLVTKFNSIMEILRPYGAFWLTAAFCILSVLF 93
 DB 438 FLVAQILFTWLCMK-FGLFLFFAFVYVMTIF 469

RESULT 4
 S25009
 monosaccharide transport protein STP4 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
 C:Accession: S25009
 R:Sauer, N.; Illig, J.; Baier, K.; Stadler, R.
 submitted to the EMBL Data Library, June 1992
 A:Description: A sink specific monosaccharid transporter from Arabidopsis thaliana.
 A:Reference number: S25009
 A:Accession: S25009
 A:Molecule type: mRNA
 A:Residues: 1-514 <SAU>
 A:Cross-references: EMBL:X66857; NID:g16523; PIDN:CAA47325.1; PID:g16524
 C:Genetics:
 A:Gene: STP4

C:Superfamily: glucose transport protein
 C:Keywords: sugar transport; transmembrane protein

Query Match 31.4%; Score 159.5; DB 2; Length 514;
 Best Local Similarity 37.8%; Pred. No. 4.4e-09;
 Matches 31; Conservative 19; Mismatches 31; Indels 1; Gaps 1;

OY 12 LAVGMCLFIAGFAVGMPIPWLMSEIFPLHIKGVATGCVLTNNMFMAFLVTKFNSIM 71
 DB 387 LIVALLCIVVAGFAWSGPLGWLVPSEIFLEIRSAQAQAINVSNMFTFLVAQLLTML 446

OY 72 ELRPYGAFWLTAAFCILSVLF 93
 DB 447 CHMK-FGLFFFAFVYVMTIF 467

RESULT 5
 S25015
 monosaccharide transport protein MST1 - common tobacco
 C:Species: Nicotiana tabacum (common tobacco)
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 24-Sep-1999
 C:Accession: S25015
 R:Sauer, N.; Stadler, R.
 submitted to the EMBL Data Library, June 1992
 A:Description: A sink specific monosaccharid transporter from tobacco.
 A:Reference number: S25015
 A:Accession: S25015
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-523 <SAU>
 A:Cross-references: EMBL:X66856; NID:g19884; PIDN:CAA47324.1; PID:g19885
 C:Superfamily: glucose transport protein

Query Match 31.4%; Score 159.5; DB 2; Length 523;
 Best Local Similarity 35.9%; Pred. No. 4.5e-09;
 Matches 33; Conservative 20; Mismatches 38; Indels 1; Gaps 1;

OY 2 PADVHLGLAWLAVGSMCLFIAGFAVGMPIPWLMSEIFPLHIKGVATGCVLTNNMFMA 61
 DB 377 PGDLPRKMYAIVVLFICIVVAGFAWSGPLGWLVPSEIFLEIRSAQSIVNVSMLEPT 436

OY 62 LVTKFNSIMEILRPYGAFWLTAAFCILSVLF 93
 DB 437 IVAQVFLTWLCMK-FGLFLFFAFVYVMTIF 467

RESULT 6
 T43176
 probable myo-inositol transport protein - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T43176
 R:Yoshioaka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
 DNA Res. 4, 363-369, 1997
 A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
 A:Reference number: Z17323; MUID:98162722
 A:Accession: T43176
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-376 <YOS>
 A:Cross-references: EMBL:D89252; NID:g1749711; PIDN:BA13913.1; PID:g1749712
 A:Experimental source: strain PR745
 C:Superfamily: maltose transport protein MAL61
 C:Keywords: transmembrane protein

Query Match 31.2%; Score 158.5; DB 2; Length 376;
 Best Local Similarity 31.6%; Pred. No. 4.1e-09;
 Matches 30; Conservative 22; Mismatches 38; Indels 5; Gaps 2;

OY 2 PADVHLGLAWLAVGSMCLFIAGFAVGMPIPWLMSEIFPLHIKGVATGCVLTNNMFMA 57

```

Db      || | | : : : : | | | | : | | : : : | |
255 PADTQNTSHSMQVWVLAIIIFLAIVASGIGNIPW-QQALEPMEYKALGAGFTAIINW 313
Oy      58 FMAFLVTKFEINSIMEILRPYGAEMLTAAFCILSYL 92
314 VGNLIISASFLPMNESITTPGTALFEGPFGVGLV 348

```

RESULT 7
E96696
Protein F1N21.12 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: F96696
R:Thellogiis, A.; Ecker, J. R.; Palm, C. J.; Federpiel, N. A.; Kaul, S.; White, O.; Alonso-
Chui, C. W.; Chung, M. K.; Comp, E.; Conway, A. B.; Conway, A. R.; Creasy, T. H.; Dewar, K.;
ansen, N. F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J. L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C. A.; Li, J. H.; Li, Y.; Lin, X.; Liu, S. X.; Liu, Z. A.; Lueros, J. S.; Maitl, R.; Marzalli,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S. L.; Schwartz, J. R.; Shinn, P.; Southwick, A. M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C. M.; Venter, J. C.; Davis, R. W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719
A:Accession: F96696
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-493 <STO>
A:Cross-references: GB:AF005173; NID:g9828628; PIDN:AAC00251.1; GSPDB:GN00141
C:Genetics:
A:Gene: F1N21.12
A:Map position: 1
A:Superfamily: glucose transport protein

RESULT 8
G84864
Probable membrane transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence-revision 02-Feb-2001 #text-change 16-Feb-2001
C:Accession: G84864
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.;
Euss, D.; Niernan, M.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-766, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487
A:Accession: G84864
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-521 <STO>
A:Cross-references: GB:AF002093; NID:g2289003; PIDN:AB64352.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g43330
A:Map position: 2
A:Superfamily: glucose transport protein

Best Local Similarity	36.1%	Pred. No. 6.4e-09.			
Matches	30;	Conservative	15;	Mismatches	38;
				Indels	0;
				Gaps	0.
QY	11	WLAAGSCLEFETIGFAVNGGPIPLMLMSELEFPIHKGATGVCVLNMFMAFLVTKKEFNSI	70		
		: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :			
Db	369	WLAVGLALAIYVFPAFGMGKPVPMYNSLEYPOYRGICGSGMATAVNMISNLVAQELTI	448		
QY	71	MEILRPYGAFFMLTAFCILSVLF	93		
		: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :			
Db	449	AEAACTGMEFLLAGIAVLAVLIF	471		

RESULT 9
SI2042
glucose transport protein STP1 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text-change 24-Sep-1999
C:Accession: SI2042; S14627
R:Sauer, N.; Friedlaender, K.; Graehl-Wilcke, U.
EMBO J. 9, 3045-3050, 1990
A:Title: Primary structure, genomic organization and heterologous expression of a glu
A:Reference number: SI2042; MUID:91005995
A:Accession: SI2042
A:Molecule type: mRNA
A:Residues: 1-522 <SAU>
A:Cross-references: EMBL:X53550; NID:q16519; PDB:CAA3037.1; PID:q16520
A:Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 459
C:Genetics:
A:Gene: STP1
C:Superfamily: glucose transport protein
C:Keywords: sugar transport; transmembrane protein

RESULT 10
E86246
glucose transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_Change 31-Mar-2001
C:Accession: E86246
R:Thelogiadis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Aloni
Chih, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huitzart, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitly, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: E86246
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-522 <ST0>
A:Cross-references: CB:AE005172; NID:95734730; PIDN:AD49955.1; GSPDB:CN00141
C:Genetics:
A:Map position: 1
A:Superfamily: glucose transport protein

RESULT 11
B96829
Probable sugar transporter, 77409-81599 [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: B96829
R:Theliositis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
Asensio, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: AB6141; MUID:21016719
A:Accession: B96829
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-467 <STO>
A:Cross-references: GB:AE005173; NID:66453876; PIDN:AA09060.1; GSPDB:GN00141
C:Genetics:
A:Gene: F19K16.22
A:Map position: 1
C:Superfamily: glucose transport protein

Query Match	30.6%	Score 155.5	DB 2:	Length 490:
Best Local Similarity	39.0%	Pred. No. 11	e-08:	
Matches	32:	Conservative	16:	Mismatches 33; Indels 1; Gaps 1;
QY	.12	LAVGSMCLFIAEAGVGGPIPELMLSEIFPLHIKGVAGVCVLTJNMFALVETKEPNSIM	71	
DB	388	LSVGVGVAANVYTFSSIGIGIPWIIINSEIPLINIKGLAGSIATLANMFAMVITVTAN-IM	446	
QY	72	ELIRPYGAFLTAAFCILSLVF	93	
DB	447	LSWNSGIFPSIYMVVCAFTVAE	468	

```

RESULT .13
S38435
hexose transport protein HUP3 - Chlorella kessleri
C:Species: Chlorella kessleri
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C/Stadler, R.; Sauer, N.
A:Accession: S38435
Submitted to the EMBL Data Library, October 1993
A:Reference number: S38435
A:Molecule type: mRNA
A:Residues: 1-534 <STA>
A:Cross-references: EMBL:X75440; NID:g408808; PTDN:CAA53192.1; PID:g408809
C:Genetics:
A:Gene: HUP3
C:Superfamily: glucose transport protein
C:Keywords: sugar transport; transmembrane protein

Query Match          30.6%; Score 155.5; DB 2; Length 534;
Best local similarity 38.5%; Pred. No. 1.2e+08;
Matches 35; Conservative 20; Mismatches 31; Indels 5; Gaps 3;

QY      2 PAIVHIGLAMLVAGSMCLPIAGAVGMPPLWLMSEIPEFLIKGVATGCVLTNNPMAF 61
        ||||| |::| ::||| ||||| :|:|||| :|: ||| :|: ||| :|: |||
Db       386 PHPSVSGV-LAA--TCIFITFAWSMGPGMWLIPEISFIETLRPAQTAVAAMGNLFSSF 441
        ::| |::| ::| |::| |::| |::| |::| |::| |::| |::| |::|

QY      62 LVTKFENSIIMELRPYGAFMLTAFCILSVL 92
Db       442 VLGQAFVSMLCAK-K-GVFLEFGAGWLVIWVL 471
        ::| |::| ::| |::| |::| |::| |::| |::| |::| |::| |::|

RESULT 14
E70070
metabolite transport protein homolog ywtG - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: E70070
R.Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berry, C.; Bron, S.; Brouillet, S.; Bruschl, C.V.; Caldwell, B.; Capriano, V.; Carter, N.M.; Auerbach, S.D.; Emerson, P.T.; Estlin, K.D.; Errington, J.; Fabret, C.; Ferrari, Nature 300, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funai, S.; Galizzi, A.; Gallardo, J.; Hardwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullio, T.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinoglu, Y.; Ogawa, K.; Ogatawa, A.; Oudega, B.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; May, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, T.; Schlegel, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Skerfving, A.; Steinhilber, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Toironi, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida, A.; Authors: Yoshikawa, H.F.; Zundstein, E.; Yoshikawa, H.; Yashochin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A:Reference number: A69580; MUID:98044033
A:Accession: E70070
A>Status: Preliminary: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-457 <KUN>
A:Cross-references: GB:A299122; GB:AL009126; NID:g2636029; PTDN:CAB15600.1; PID:g26361
```


A:Experimental source: strain 168
 C:Genetics:
 A:Gene: YWtG
 C:Superfamily: glucose transport protein

Query Match 30.5%; Score 155; DB 2; Length 457;
 Best Local Similarity 35.7%; Pred. No. 1.2e-08;
 Matches 30; Conservative 19; Mismatches 35; Indels 0; Gaps 0;

OY 10 AMLAAGSMCLFIAGFVAGCPILPMLMSEIFPHIKGVATGVCVLTNWMFMAFLVTKFEFS 69

DB 336 SWTVYICLVFIVFAVSWGCPVWMLPELPELHVRGIGTGVSTLMLHVGTLIVSLTPPT 395

OY 70 IMEILRPYGAFMLTAFCILSVLF 93

DB 396 LMEALICISYLFILYAAIGIMAFLF 419

RESULT 15

D96589

hypothetical protein T22H22.15 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: D96589

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: D96589

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-483 <STO>

A:Cross-references: GB:AE005173; NID:g3776581; PIDN:AAC64898.1; GSPDB:GN00141

C:Genetics:

A:Gene: T22H22.15

A:Map position: 1

Query Match 30.4%; Score 154.5; DB 2; Length 483;

Best Local Similarity 33.3%; Pred. No. 1.4e-08;

Matches 31; Conservative 24; Mismatches 37; Indels 1; Gaps 1;

OY 1 EPADVHLGLAMLAAGSMCLFIAGFVAGCPILPMLMSEIFPHIKGVATGVCVLTNWMFMA 60

DB 367 ESAGVSSKIGMILAMVYVYTGSFSLMGCTIPVIMSELFPIIDIKSAGSLTVVSWGS 426

OY 61 FLVTKFEFSIMEILRPYGAFMLTAFCILSVLF 93

DB 427 WLISFTFNLMN-WNPAGTFYVATVCGATVIF 458

Search completed: February 13, 2002, 21:53:11
 Job time: 6133 sec

Matches	34	Conservative	17	Mismatches	40	Indels	1	Gaps	1
Oy	2	PADVHLGLAMLVAGSMCLFIAGFANGKQPIPLMLSEIFPLHIKQVAGVCVLTWPMFAP	61						
Db	381	PDGIDPQWVAVVVVLFICITIVYSGFAMSGMPLGLVLPSEIFPLEIRSAOSVNVSVMMFTTF	440						
Oy	62	LVTFEFSIMETLBPYGAFWLTAAFCIISVLF	93						
Db	441	VVAQVF-LIMLCHKLFGFLFFSFVFLVMSIF	471						
RESULT	4								
AC	HUPL_CHLKE	STANDARD	PRT: 534 AA.						
AC	P15686; Q39523;								
DT	01-APR-1998 (Rel. 14, Created)								
DT	15-JUL-1998 (Rel. 36, Last sequence update)								
DT	15-JUL-1998 (Rel. 36, Last annotation update)								
DE	H(+)/HEXOSE COTRANSPORTER 1.								
GN	HUPL.								
OS	Chlorella kessleri.								
OX	Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorella.								
OX	NCBI_TaxID=3074;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=90092536; PubMed=2599110;								
RA	Sauer N., Tanner W.;								
RT	"the hexose carrier from Chlorella. cDNA cloning of a eucaryotic H+-								
RT	cotransporter.";								
RL	FEBS Lett. 259:43-46(1989).								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=91330335; PubMed=1868571;								
RA	Wolf K., Tanner W., Sauer N.;								
RT	"the Chlorella H+/hexose cotransporter gene.";								
RL	Curr. Genet. 19:215-219(1991)								
CC	-1- FUNCTION: ACTIVE UPTAKE OF HEXOSES.								
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.								
CC	-1- INDUCTION: BY GLUCOSE.								
CC	-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/								
CC	or send an email to license@isb-sib.ch)								
CC	-----								
DR	EMBL; Y07520; CAA68813.1; -;								
DR	EMBL; X55349; CAA39036.1; -;								
DR	PIR; S07096; S07096.								
DR	InterPro; IPRO003663; Sugar_tnsportr.								
DR	InterPro; IPRO003662; sub_tnsportr.								
DR	Pfam; PF00083; sugar_tr.1.								
DR	PRINTS; PR00171; SDGRTNSPORT.								
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.								
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.								
KW	Transmembrane; Transport; Sugar transport; Symport; Repeat;								
KW	Multigene family.								
FT	TRANSMEM	26	46	POTENTIAL.					
FT	TRANSMEM	88	108	POTENTIAL.					
FT	TRANSMEM	117	137	POTENTIAL.					
FT	TRANSMEM	145	165	POTENTIAL.					
FT	TRANSMEM	172	192	POTENTIAL.					
FT	TRANSMEM	202	222	POTENTIAL.					
FT	TRANSMEM	288	308	POTENTIAL.					
FT	TRANSMEM	323	344	POTENTIAL.					
FT	TRANSMEM	353	373	POTENTIAL.					
FT	TRANSMEM	388	408	POTENTIAL.					
FT	TRANSMEM	427	447	POTENTIAL.					
FT	TRANSMEM	457	477						

```

FT      CONFLICT      209      209      P -> L (IN REF. 1)).
SO      CONFLECT      248      248      C -> R (IN REF. 1)).
FO      SEQUENCE      534 AA; 57522 MW; 3B238B8710AA2237 CRC64;

Query Match          30.0%; Score 152.5; DB 1; Length 534;
Best Local Similarity 38.5%; Pred. No. 2.9e-07;
Matches   35; Conservative 19; Mismatches 32; Indels 5; Gaps 3;

QY      2 PADVHLGAMLAVGSMLCFIAGFAVGCGPIPLMLSEIFPHLIKAVATGCVLTNMFMAF 61
       1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      386 PKRNASGI--LAV--ICLFISGFASMSWPMGLMSELFLETETPRAGRAVAAVGVNLELSF 441
       ::::: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY      62 LVTKFFNSIMELRLPYGAFWLTAFACILSLV 92
       442 VIGQAFAVSMLCAME-YGVLEFFAGLVIMVL 471

RESULT      5
ITR2_SCHPO ITR2_SCHPO STANDARD; PRT; 557 AA.
AC      P87110; P78901;
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      MYO-INOSTIOL TRANSPORTER 2.
GN      ITR2 OR SPAC20G8.03.
OS      Schizosaccharomyces pombe (Fission yeast).
OC      Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
        Schizosaccharomycetales; Schizosaccharomycetaceae;
        Schizosaccharomycetes.
OX      NCBI_TaxID=4896;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=968 H90;
RX      MEDLINE=96828265; PubMed=9560432;
RA      Niederberger C., Graub R., Schweinrubner A.-M., Fankhauser H.,
RA      Rusu M., Poitelea M., Edenharter L., Schweinrubner M.E.;
RT      "Exogenous Inositol and genes responsible for inositol transport are
RT      required for mating and sporulation in Shizosaccharomyces pombe.";
RT      Curr. Genet. 33:255-261(1998).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=972;
RA      Badcock K., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;
RL      Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE OF 166-541 FROM N.A.
RC      STRAIN=PR745;
RA      Yoshioaka S., Kato K., Okayama H.;
RL      Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: TRANSPORTER FOR MYO-INOSTIOL.
CC      -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC      -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outpost at
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL: X99105; -: NOT ANNOTATED_CDS.
DR      EMBL: Z95334; CAB08597.1; -.
DR      EMBL: D89252; BAAL3913.1; -.
DR      InterPro: IPRO03663; Sugar_tnsportr.
DR      InterPro: IPRO03662; sub_tnsportr.
DR      Pfam: PF00083; sugar_tf_1.
DR      PRINTS: PR00171; SUGRTNSPORT.
DR      PRINTS: PR00172; GLUCTRNSPORT.
DR      ProSITE: PS00216; SUGAR_TRANSPORT_1; 2.
DR      ProSITE: PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
```

KW Transmembrane; Sugar transport; Glycoprotein.
 FT TRANSMEM 77 97 POTENTIAL.
 FT TRANSMEM 100 120 POTENTIAL.
 FT TRANSMEM 124 144 POTENTIAL.
 FT TRANSMEM 158 178 POTENTIAL.
 FT TRANSMEM 181 201 POTENTIAL.
 FT TRANSMEM 210 230 POTENTIAL.
 FT TRANSMEM 241 261 POTENTIAL.
 FT TRANSMEM 368 388 POTENTIAL.
 FT TRANSMEM 397 417 POTENTIAL.
 FT TRANSMEM 433 453 POTENTIAL.
 FT TRANSMEM 469 489 POTENTIAL.
 FT TRANSMEM 499 519 POTENTIAL.
 FT DOMAIN 520 557 CYTOPLASMIC (POTENTIAL).
 FT CARBOHD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 271 271 N -> H (IN STRAIN PR745).
 FT VARIANT 428 428 N -> H (IN STRAIN PR745).
 FT VARIANT 436 436 L -> I (IN STRAIN PR745).
 FT VARIANT 526 526 S -> F (IN STRAIN PR745).
 SQ SEQUENCE 557 AA; 61136 MW; 20875EC11B153175 CRC64;

Query Match 29.4%; Score 149.5; DB 1; Length 557;
 Best Local Similarity 30.5%; Pred. No. 5.6e-07;
 Matches 29; Conservative 23; Mismatches 38; Indels 5; Gaps 2;

OY 2 PAD---VHLGLAVGSMCLFIAGFVAGWGPITPMLLSEIFPHIKGVATGCVLTNW 57
 DB 420 PADTTQNTNSGQYVYVLAIIIFLASYSAGIGNITP-QQAEIFPHEVRALGAGFSTAINW 478
 OY 58 FMAFLVTKFNSIMEILRPYGAFWLTAAFCILSVL 92
 DB 479 VGNLIISASFLTMESTPTPTGFALFAGCFVGLV 513

RESULT 6
 ARAE_KLEOX STANDARD; PRT; 472 AA.
 ID ARAE_KLEOX
 AC P45598;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 36, Last annotation update)
 DE ARABINOSE-PROTON SYMPORTER (ARABINOSE TRANSPORTER).
 GN ARAE.
 OS Klebsiella oxytoca.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Klebsiella.
 OX NCBI_TaxID=571;
 OX [1]
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=8017;
 RX MEDLINE=95394866; PubMed=7665532;
 RA Shitwell K.P., Charalambous B.M., McDonald T.P., Henderson P.J.F.;
 RT "Cloning, sequencing, and expression of the araE gene of Klebsiella
 oxytoca 8017, which encodes arabinose-H⁺ symport activity.";
 RL J. Bacteriol. 177:5379-5380(1995).
 CC -1- FUNCTION: UPTAKE OF ARABINOSE ACROSS THE BOUNDARY MEMBRANE WITH
 THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X79598; GA56110.1; -
 DR InterPro: IPR003663; Sugar_trnsport.
 DR InterPro: IPR003662; sub_trnsport.
 DR Pfam: PF00083; sugar_trf.1.

DR PRINTS; PR00171; SUGRTNSPORT.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport.
 FT TRANSMEM 23 43 POTENTIAL.
 FT TRANSMEM 63 83 POTENTIAL.
 FT TRANSMEM 92 110 POTENTIAL.
 FT TRANSMEM 120 138 POTENTIAL.
 FT TRANSMEM 149 167 POTENTIAL.
 FT TRANSMEM 177 197 POTENTIAL.
 FT TRANSMEM 259 278 POTENTIAL.
 FT TRANSMEM 299 318 POTENTIAL.
 FT TRANSMEM 328 348 POTENTIAL.
 FT TRANSMEM 362 382 POTENTIAL.
 FT TRANSMEM 396 415 POTENTIAL.
 FT TRANSMEM 424 443 POTENTIAL.
 SQ SEQUENCE 472 AA; 51732 MW; 410021E1BF3D96E CRC64;

Query Match 28.8%; Score 146.5; DB 1; Length 472;
 Best Local Similarity 34.9%; Pred. No. 9.3e-07;
 Matches 30; Conservative 16; Mismatches 37; Indels 3; Gaps 1;

OY 8 GLAVLAVGSMCLFIAGFVAGWGPITPMLLSEIFPHIKGVATGCVLTNFWAFLVTKF 67
 DB 356 GLSWLSVGTWMCIGAGYMASAPVWIIQSEIQPLKCRDFGTCSTTNWVSNMIGATF 415
 OY 68 NSIMEILRPYGAFWLTAAFCILSVL 93
 DB 416 LTLDDAIGAGTFWLYTA---LNVAF 438

RESULT 7
 ARAE_ECOLI STANDARD; PRT; 472 AA.
 ID ARAE_ECOLI
 AC P09630; Q46937;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE ARABINOSE-PROTON SYMPORTER (ARABINOSE TRANSPORTER).
 GN ARAE OR B2841 OR Z4161 OR ECS3698.
 OS Escherichia coli, and
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 OX NCBI_TaxID=562, 83334;
 OX [1]
 RN RP SEQUENCE FROM N.A.
 RC MEDLINE=87115869; PubMed=3543693;
 RA Maiden M.C.J., Davis E.O., Baldwin S.A., Moore D.C.M.,
 RA Henderson P.J.F.;
 RT "Mammalian and bacterial sugar transport proteins are homologous.";
 RL Nature 325:641-643(1987).
 CC [2]
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / JM2433;
 RX MEDLINE=88228015; PubMed=2836407;
 RA Maiden M.C.J., Jones-Mortimer M.C., Henderson P.J.F.;
 RT "The cloning, DNA sequence, and overexpression of the gene araE
 coding for arabinose-proton symport in Escherichia coli K12.";
 RL J. Biol. Chem. 263:8003-8010(1988).
 CC [3]
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1235-1244(1997).
 CC [4]
 RN RP SEQUENCE FROM N.A.

RC STRAIN-0157:H7 / EDL933 / ATCC 700927;
 RA MEDLINE-21074935; PubMed-11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouzis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-0157:H7 / RIMD 0509952;
 RA MEDLINE-21156231; PubMed-11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 RN [6]
 RP PRELIMINARY SEQUENCE OF 1-28 FROM N.A.
 RA MEDLINE-84114868; PubMed-6319708;
 RA Stoner C., Schleif R.F.;
 RT "The arae low affinity L-arabinose transport promoter. Cloning,
 RT sequence, transcription start site and DNA binding sites of
 RT regulatory proteins.";
 RL J. Mol. Biol. 171:369-381(1983).
 CC -i- FUNCTION: UPTAKE OF ARABINOSE ACROSS THE BOUNDARY MEMBRANE WITH
 CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
 CC -i- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC -i- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: J03732; AAA23469.1; -;
 DR EMBL: X00272; CAA25075.1; -;
 DR EMBL: U29581; AAB40488.1; -;
 DR EMBL: AE000368; AAC75880.1; -;
 DR EMBL: AE005513; AAG57953.1; -;
 DR EMBL: AP002563; BAB37121.1; -;
 DR PIR: B26430; B26430.
 DR PIR: A28075; A28075.
 DR Ecogene: EG10056; arae.
 DR InterPro: IPR003663; Sugar_tnsportr.
 DR InterPro: IPR003662; sub_tnsportr.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGRTNSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transport: Sugar transport; Transmembrane; Inner membrane; Symport;
 KW Complete proteome.
 FT TRANSMEM 23 43 POTENTIAL.
 FT TRANSMEM 63 83 POTENTIAL.
 FT TRANSMEM 92 110 POTENTIAL.
 FT TRANSMEM 120 138 POTENTIAL.
 FT TRANSMEM 149 167 POTENTIAL.
 FT TRANSMEM 177 197 POTENTIAL.
 FT TRANSMEM 259 278 POTENTIAL.
 FT TRANSMEM 299 318 POTENTIAL.
 FT TRANSMEM 328 348 POTENTIAL.
 FT TRANSMEM 362 382 POTENTIAL.
 FT TRANSMEM 396 415 POTENTIAL.
 FT TRANSMEM 424 443 POTENTIAL.
 FT TRANSMEM 426 28 POTENTIAL.
 FT CONFILIT 26 28 SVA -> YDR (IN REF. 6).
 SQ SEQUENCE 472 AA: 51684 MW: 411904441D4393 CRC64:

Query Match 28.7%; Score 146; DB 1; Length 472;
 Best Local Similarity 34.5%; Pred. No. 1e-06;
 Matches 30; Conservative 15; Mismatches 42; Indels 0; Gaps 0;
 OY 8 GIAMLAAGSMCLFIACFANGKPIPLMLSEIFPLHIKAVGVCYLTWFFAFLVTKRF 67
 DB 356 GLSWLSVGTMTMCICAGYAMSAAPVWILCSEIPLKCRDFICSTTTWVSNMIIGATF 415
 OY 68 NSIMELIRPYGAFWLTAAFCILSVLET 94
 DB 416 LTLIDSIGAGTFWLTALNIAFVGIT 442
 RESULT 8
 HEX6_RICCO STANDARD: PRT; 510 AA.
 AC 007423;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE HEXOSE CARRIER PROTEIN HEX6.
 GN HEX6.
 OS Ricinus communis (Castor bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.
 OX NCBI_TaxID=3988;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. CARMENCITA; TISSUE=COyledon;
 RA Weid A., Franz J., Sauer N., Komor E.;
 RT "Isolation of a family of cDNA-clones from Ricinus communis L.
 RT with close homology to the hexose carriers.";
 RL J. Plant Physiol. 143:178-183(1994).
 CC -i- FUNCTION: ACTIVE UPTAKE OF HEXOSES. PROBABLE GLUCOSE/HYDROGEN
 CC SYMPORT.
 CC -i- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -i- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: I08188; AAA79857.1; -;
 DR InterPro: IPR003663; Sugar_tnsportr.
 DR InterPro: IPR003662; sub_tnsportr.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGRTNSPORT.
 DR PROSITE: PS00172; GLUCTNSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane; Transport; Sugar transport; Symport.
 FT TRANSMEM 21 41 POTENTIAL.
 FT TRANSMEM 80 100 POTENTIAL.
 FT TRANSMEM 118 138 POTENTIAL.
 FT TRANSMEM 140 160 POTENTIAL.
 FT TRANSMEM 167 187 POTENTIAL.
 FT TRANSMEM 202 222 POTENTIAL.
 FT TRANSMEM 284 304 POTENTIAL.
 FT TRANSMEM 322 342 POTENTIAL.
 FT TRANSMEM 349 369 POTENTIAL.
 FT TRANSMEM 386 406 POTENTIAL.
 FT TRANSMEM 425 445 POTENTIAL.
 FT TRANSMEM 451 471 POTENTIAL.
 SQ SEQUENCE 510 AA: 55594 MW: 26615F8EBD84299E CRC64:

```

Query Match Similarity      27.9%; Score 141.5; DB 1; Length 510;
Best Local Similarity      30.2%; Pred. No. 2.9e-06;
Matches 26; Conservative 22; Mismatches 37; Indels 1; Gaps 1;

QY      8 GLAWLVGSMCLFTIAGFAVWGPIPLMLSEFFPLHIKVGATGVCVLTNMFAFLVTKEF 67
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      382 GYAVITLLICITIVYAGFCMSWGLCLMVPSEIFPLEISAGSIVVAVSFLFTFVAQTF 441
QY      68 NSIMEILRPYCAFWLTLTAFCILSYLVE 93
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      442 LSMLCHEKS-GIIEFFGGWVYVMTAF 466

RESULT      9
GALP_ECOLI
AC      P37021; STANDARD; PRT; 464 AA.
DT      01-JUN-1994 (Rel. 29, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DT      20-AUG-2001 (Rel. 40, last annotation update)
DE      GALACTOSE-PROTON SYMPORTER (GALACTOSE TRANSPORTER).
GN      GALP OR B2943.
OC      Escherichia coli.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
OX      NCBI_TaxID=562;
RN      [1]
RS      SEQUENCE FROM N.A., AND CHARACTERIZATION.
RA      Roberts P.E.;
RL      Thesis (1992), University of Cambridge, U.K.
RP      [2]
RX      STRAIN-K12 / MG1655.
RX      MEDLINE=9742617; PubMed=9278503;
RA      Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Shao Y.;
RT      "The complete genome sequence of Escherichia coli K-12.";
RL      Science 277:1453-1474(1997).
CC      -1- FUNCTION: UPTAKE OF GALACTOSE ACROSS THE BOUNDARY MEMBRANE WITH
CC      THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC      -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
CC      -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: U28377; AAA69110.1; -.
DR      EMBL: AE000377; AAC75980.1; -.
DR      EcoGene: EG12148; galP.
DR      InterPro: IPR003663; Sugar_tnsport.
DR      InterPro: IPR003662; sub_tnsport.
DR      Pfam: PF00083; sugar_tr.1.
DR      PRINTS: PR00171; SUGRTNSPORT.
DR      PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR      PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW      Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
KW      Complete proteome.
FT      TRANSMEM      16      36      POTENTIAL.
FT      TRANSMEM      57      77      POTENTIAL.
FT      TRANSMEM      85      105      POTENTIAL.
FT      TRANSMEM      113     133      POTENTIAL.
FT      TRANSMEM      140     160      POTENTIAL.
FT      TRANSMEM      172     192      POTENTIAL.
FT      TRANSMEM      251     271      POTENTIAL.
FT      TRANSMEM      291     311      POTENTIAL.
FT      TRANSMEM      322     342      POTENTIAL.

```

FT	TRANSMEM	352	372	POTENTIAL.
FT	TRANSMEM	395	415	POTENTIAL.
FT	TRANSMEM	417	437	POTENTIAL.
SO	SEQUENCE	464 AA;	50982 MW;	07E08935BD8E3F8E CRC64;
Query Match 27.7%; Score 140.5; DB 1; Length 464;				
Best Local Similarity 31.6%; Pred No. 3.3e-06;				
Matches 30; Conservative 18; Mismatches 42; Indels 5; Gaps 1.				
QY	5 VHIG-----MLAAGSMCLFIAGFAVAGCPILPMLMSEIFPLHIKGVATGVCLTNMEM 59			
DB	338 MHIGHSASQAFALAMLMFLTVGFAMSGPLIWLICSGIDPLKGRDFOITGSTATNMIA 397			
QY	60 AFLVTKFNSIMELILPYGAFLWIAFCILSLVLT 94			
DB	398 NMIVGATFLTMLNTLGNANTFVWYALANVLFIILT 432			
RESULT 10				
ID	XYLE_ECOLI	STANDARD:	PRT:	491 AA.
AC	P09098:			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-MAR-1989 (Rel. 10, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	D-XYLOSE-PROTON SYMPORTER (D-XYLOSE TRANSPORTER).			
GN	XYLE OR B4031 OR Z5629 OR ECS5014.			
OS	Escherichia coli, and			
OS	Escherichia coli O157:H7.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Escherichia.			
OX	NCBI_TaxID=562, 83334;			
LN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12;			
RX	MEDLINE=88007632; PubMed=2820984;			
RT	Davis E.O., Henderson P.J.F.;			
RT	"The cloning and DNA sequence of the gene xyle for xylose-proton			
RT	symport in Escherichia coli K12.";			
RL	J. Biol. Chem. 262:13928-13932(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87115869; PubMed=3543693;			
RA	Maiden M.C.J., Davis E.O., Baldwin S.A., Moore D.C.M.,			
RA	Henderson P.J.F.;			
RT	"Mammalian and bacterial sugar transport proteins are homologous.";			
RL	Nature 325:641-643(1987).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12 / MG1655;			
RX	MEDLINE=94089392; PubMed=8265357;			
RA	Blattner F.R., Burland V.D., Plunkett G. II, Sofia H.J.,			
RA	Daniels D.L.;			
RT	"Analysis of the Escherichia coli genome. IV. DNA sequence of the			
RT	region from 89.2 to 92.8 minutes.";			
RL	Nucleic Acids Res. 21:5408-5417(1993).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=O157:H7 / EDL933 / ATCC 700927;			
RX	MEDLINE=21074935; PubMed=11206551;			
RA	Perna N.T., Plunkett G. II, Burland V., Mau B., Glasner J.D.,			
RA	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,			
RA	Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,			
RA	Grothbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,			
RA	Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,			
RA	Welch R.A., Blattner F.R.;			
RT	"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";			
RL	Nature 409:529-533(2001).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=O157:H7 / RIMD 0509952;			
RX	MEDLINE=21156231; PubMed=11258796;			

RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.,
 RT "Complete genome sequence of enterohemorrhagic *Escherichia coli*
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RT DNA Res. 8:11-22(2001).
 RN [6]
 RP SEQUENCE OF 1-192 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=88234001; PubMed=2836810;
 RA Francoz E., Dassa E.;
 RT "3' end of the maleFG operon in *E. coli*: localization of the
 RT transcription termination site.";
 RT Nucleic Acids Res. 16:4097-4109(1988).
 CC -1- FUNCTION: UPTAKE OF D-XYLOSE ACROSS THE BOUNDARY MEMBRANE WITH
 CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC -1- INDUCTION: BY XYLOSE.
 CC -1- MISCELLANEOUS: *E. COLI* HAS TWO D-XYLOSE TRANSPORT SYSTEMS THAT
 CC ACCUMULATE SUGAR AGAINST A CONCENTRATION GRADIENT: THE XYLE SYSTEM
 CC WHICH UTILIZES THE ELECTROCHEMICAL GRADIENT OF PROTONS AND THAT IS
 CC INSENSITIVE TO COLD OSMOTIC SHOCK AND THE XYLEP SYSTEM THAT USES A
 CC HIGH-ENERGY PHOSPHATE COMPOUND AND IS SENSITIVE TO COLD OSMOTIC
 CC SHOCK.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J02812; AAA79016.1; -;
 DR EMBL: U00006; AAC3125.1; -;
 DR EMBL: AE000476; AAC77001.1; -;
 DR EMBL: AE005636; AAG59230.1; -;
 DR EMBL: AP002568; BAB38437.1; -;
 DR EMBL: X06663; CAA29863.1; -;
 DR PIR: A26430; A26430;
 DR PIR: A27418; A27418;
 DR PIR: S00874; S00874;
 DR EcoGene: EGI1076; xyle.
 DR InterPro: IPR003663; Sugar_tnsport.
 DR InterPro: IPR003662; sub_tnsport.
 DR Pfam: PF00083; sugar_trf.1.
 DR PRINTS: PR00171; SUGRTNSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
 KW Complete proteome.
 FT TRANSMEM 12 32 POTENTIAL.
 FT TRANSMEM 57 77 POTENTIAL.
 FT TRANSMEM 91 111 POTENTIAL.
 FT TRANSMEM 135 155 POTENTIAL.
 FT TRANSMEM 167 187 POTENTIAL.
 FT TRANSMEM 202 222 POTENTIAL.
 FT TRANSMEM 274 294 POTENTIAL.
 FT TRANSMEM 314 334 POTENTIAL.
 FT TRANSMEM 344 364 POTENTIAL.
 FT TRANSMEM 371 391 POTENTIAL.
 FT TRANSMEM 409 429 POTENTIAL.
 FT TRANSMEM 445 465 POTENTIAL.
 FT CONFLICT 64 64 A -> V (IN REF. 6).
 SO SEQUENCE 491 AA; 53608 MW; 2AF1AF9756C0B722 CRC64;

Query Match 27.7%; Score 140.5; DB 1; Length 491;
 Best Local Similarity 33.0%; Pred. No. 3, 5e-06;
 Matches 33; Conservative 19; Mismatches 35; Indels 13; Gaps 3;

OY 7 LGLAW-----LAVGSCLEIFAGVWGPIPWILNSELPLHIKVGATGVLTNFM 59
 DB 359 LGTAFFYTOPAPICVALLSMIFYAAFGMSGVCVLLSEIFPNAIRKALAAVAOWLA 418
 OY 60 AFLVTKER-----NSIMELRLPYG-AFWLTAFCFLSLVF 93
 DB 419 NYFVSWTEPPMKDKNSLVAFHFNHGSFYWIYGCMGVLAALF 458
 RESULT 11
 CSBC_BACSU STANDARD; PRT: 461 AA.
 AC P46333; 032289;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROBABLE METABOLITE TRANSPORT PROTEIN CSBC.
 GN CSBC OR SS92BR.
 OS *Bacillus subtilis*.
 OC Bacteria; Firmicutes; *Bacillus*/Clostridium group;
 CC *Bacillus*/Staphylococcus group; *Bacillus*.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / BGSC1A1;
 RX MEDLINE=96093926; PubMed=7584049;
 RA Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
 RT "Cloning and sequencing of a 36-kb region of the *Bacillus subtilis*
 RT genome between the *gnt* and *iol* operons.";
 RT DNA Res. 2:61-69(1995).
 RN [2]
 RP REVISIONS.
 RA Fujita Y., Shibayama T., Ishio I., Aoyama D., Yoshida K.-I.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=99303315; PubMed=10376822;
 RA Akbar S., Lee S.-Y., Boylan S.A., Price C.W.;
 RT "Two genes from *Bacillus subtilis* under the sole control of the
 RT general stress transcription factor *sigmas*.";
 RL Microbiology 145:1069-1078(1999).
 CC -1- FUNCTION: COULD SERVE EITHER A NUTRITIONAL OR AN OSMOTIC
 CC PROTECTION FUNCTION.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB005554; BAA21604.1; -;
 DR EMBL: Z99124; CAB16017.1; -;
 DR Subtilist; BG11360; csbc.
 DR InterPro: IPR003663; Sugar_tnsport.
 DR InterPro: IPR003662; sub_tnsport.
 DR Pfam: PF00083; sugar_trf.1.
 DR PRINTS: PR00171; SUGRTNSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 2.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 15 35 POTENTIAL.
 FT TRANSMEM 39 59 POTENTIAL.
 FT TRANSMEM 77 97 POTENTIAL.
 FT TRANSMEM 105 125 POTENTIAL.

Query Match	Best Local Similarity	Score	DB 1;	Length	575;
Matches	30; Conservative	15;	Mismatches	32;	Indels
				13;	Gaps
					3;
Qy	10 ANIANG--SMCLFICFANGGPIPLWLMSELPFLHIGVATGCVLNNFPAFLVTEEF	67			
Db	440 AWAIVLISMIYVSVSYASGLGNLFM--OOSLELFPMSVRGLGTGMSTAVVMAGNLIGASF	498			
Qy	68 NSIMETILRPYGA-----WFLTAFC	87			
Db	499 LTLMSFTPTTGTFALYGLGCLFLMGALFC	528			
RESULT 13					
GTR2_MOUSE					
AC	PIA246;	STANDARD:	PRT:	523 AA.	
DT	01-JAN-1990 (Rel. 13, Created)				
DT	01-FEB-1996 (Rel. 33, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	GLUCOSE TRANSPORTER TYPE 2, LIVER.				
GN	SLC2A2 OR GLUT2 OR GLUT-2.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_Taxid=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=BALB/C; TISSUE=Liver;				
RC	STRAIN=C57BL/6; TISSUE=Liver;				
RC	MEDLINE=90098776; PubMed=2602116;				
RA	Suzue K., Lodish H.F., Thorens B.;				
RT	"Sequence of the mouse liver glucose transporter.";				
RL	Nucleic Acids Res. 17:10099-10099(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6; TISSUE=Liver;				
RC	MEDLINE=89366666; PubMed=2771649;				
RA	Asano T., Shibaaki Y., Lin J.L., Akanuma Y., Takaku F., Oka Y.;				
RT	"The nucleotide sequence of cDNA for a mouse liver-type glucose				
RT	transporter protein.";				
RL	Nucleic Acids Res. 17:6386-6386(1989).				
RN	[3]				
RP	SEQUENCE OF 384-496 FROM N.A.				
RC	MEDLINE=92111400; PubMed=1765007;				
RA	Hogan A., Heyner S., Charron M.J., Copeland N.G., Gilbert D.J.,				
RA	Jenkins N.A., Thorens B., Schultz G.A.;				
RT	"Glucose transporter gene expression in early mouse embryos.";				
RL	Development 113:363-372(1991).				
RN	[4]				
RP	TISSUE SPECIFICITY.				
RC	STRAIN=C57BL/6;				
RC	MEDLINE=93170163; PubMed=1289053;				
RA	Smith D.E., Gridley T.;				

RT "Differential screening of a PCR-generated mouse embryo cDNA library:
 RT glucose transporters are differentially expressed in early
 RT postimplantation mouse embryos.";
 RL Development 116:555-561(1992).
 CC -I- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY
 CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
 CC MEMBRANE OF HEPATOCYTES & IS RESPONSIBLE FOR UPTAKE OF GLUCOSE BY
 CC THE BETA CELLS. IT MAY COMPRISE PART OF THE GLUCOSE-SENSING
 CC MECHANISM OF THE BETA CELL. IT MAY ALSO PARTICIPATE WITH THE
 CC NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCELLULAR TRANSPORT OF
 CC GLUCOSE IN THE SMALL INTESTINE & KIDNEY.
 CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -I- TISSUE SPECIFICITY: IN EMBRYO, EXPRESSED IN ENDODERM LAYER OF
 CC YOLK SAC AND LIVER PRIMORDIUM.
 CC -I- SIMILARITY: HIGH TO OTHER GLUCOSE TRANSPORTERS.
 CC -I- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X16986; CAA34855.1; -;
 DR EMBL: X15684; CAA33719.1; -;
 DR EMBL: S77926; AAB20847.1; -;
 DR PIR: S06920; S06920.
 DR PIR: S05319; S05319.
 DR MGD: MGI:1095438; Slc2a2.
 DR InterPro: IPR003663; Sugar_trnsport.
 DR InterPro: IPR003662; sub_trnsport.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PRO0171; SUGRTNSPORT.
 DR PRINTS: PRO0172; GLUCTRNSPORT.
 DR PRINTS: PRO1191; GLUCTRNSPORT2.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 DR Trnsmembrane: Sugar transport; Transport; Glycoprotein;
 KW Multigene family.
 FT TRANSMEM 11 31 POTENTIAL.
 FT DOMAIN 32 97 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 98 118 POTENTIAL.
 FT TRANSMEM 127 147 POTENTIAL.
 FT TRANSMEM 158 178 POTENTIAL.
 FT TRANSMEM 187 207 POTENTIAL.
 FT TRANSMEM 217 237 POTENTIAL.
 FT DOMAIN 238 302 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 303 323 POTENTIAL.
 FT TRANSMEM 338 358 POTENTIAL.
 FT TRANSMEM 369 389 POTENTIAL.
 FT TRANSMEM 400 420 POTENTIAL.
 FT TRANSMEM 433 453 POTENTIAL.
 FT TRANSMEM 461 481 POTENTIAL.
 FT DOMAIN 482 523 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 106 106 G -> D (IN REF. 2).
 FT CONFLICT 181 181 N -> T (IN REF. 2).
 FT CONFLICT 327 327 S -> T (IN REF. 2).
 FT CONFLICT 351 351 L -> F (IN REF. 2).
 FT CONFLICT 432 432 P -> S (IN REF. 1).
 SO SEQUENCE 523 AA; 57085 MW; 383F780BD73F0DFC CRC64;

Query Match 27.0%; Score 137; DB 1; Length 523;
 Best Local Similarity 34.1%; Pred. No. 7.7e-06;
 Matches 30; Conservative 17; Mismatches 37; Indels 4; Gaps 2;

OY 10 AWLAGSMC---LFTAGFVWGPIPWLMSEIFPLHKGATGVCVLTNMFMAFLVTK 66
 DB 393 AMMSVSMATLFLFVSPFELGPIPLMFWAVAFEFSGRPTALALAFSNMVCNEVIALC 452

OY 67 ENSIMEILRPYGAWLTAACILSLVFT 94
 DB 453 FOYIADFLGPY-VFLEFAGVAVLFTLFT 479

RESULT 14
 HUP2_CHLKE STANDARD: PRT; 540 AA.
 ID HUP2_CHLKE
 AC 039524;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE H(+)/HEXOSE COTRANSPORTER 2 (GALACTOSE-H+ SYMPORTER).
 GN HUP2.
 OS Chlorella kessleri.
 OC Eukaryota: Viridiplantae: Chlorophyta: Trebouxiophyceae: Chlorella.
 OX NCBI_TaxID=3074;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95175627; PubMed=7870840;
 RA Stadler R., Wolf K., Hilgath C., Tanner W., Sauer N.K.;
 RT "Subcellular localization of the inducible Chlorella HUP1
 RT monosaccharide-H+ symporter and cloning of a Co-induced galactose-H+
 RT symporter.";
 RL Plant Physiol. 107:33-41(1995).
 CC -I- FUNCTION: ACTIVE UPTAKE OF GALACTOSE.
 CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -I- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X66855; CAA47323.1; -;
 DR InterPro: IPR003663; Sugar_trnsport.
 DR InterPro: IPR003662; sub_trnsport.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PRO0171; SUGRTNSPORT.
 DR PRINTS: PRO0172; GLUCTRNSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 DR Trnsmembrane: Transport; Sugar transport; Symport; Repeat;
 KW Multigene family.
 FT TRANSMEM 29 49 POTENTIAL.
 FT TRANSMEM 91 111 POTENTIAL.
 FT TRANSMEM 125 145 POTENTIAL.
 FT TRANSMEM 148 168 POTENTIAL.
 FT TRANSMEM 175 195 POTENTIAL.
 FT TRANSMEM 208 228 POTENTIAL.
 FT TRANSMEM 297 317 POTENTIAL.
 FT TRANSMEM 327 347 POTENTIAL.
 FT TRANSMEM 358 378 POTENTIAL.
 FT TRANSMEM 397 417 POTENTIAL.
 FT TRANSMEM 433 453 POTENTIAL.
 FT TRANSMEM 459 479 POTENTIAL.
 SO SEQUENCE 540 AA; 58343 MW; 0A3801A33E0BA733 CRC64;

Query Match 26.9%; Score 136.5; DB 1; Length 540;
 Best Local Similarity 32.5%; Pred. No. 8.8e-06;
 Matches 25; Conservative 20; Mismatches 31; Indels 1; Gaps 1;

OY 17 MCLFTAGFVWGPIPWLMSEIFPLHKGATGVCVLTNMFMAFLVTKENSIHELRP 76
 DB 399 ICVVYAAAFAMSWGPIGLMVLPSSEIQTLFTFGAGMSMAVIVNLFSEVIGCAFISMCMAMR- 457

OY 77 VGAFMTAFACTLSVLF 93
 DB 458 WGVFLFPAGWVAVIMTF 474

RESULT 15
GTR2_RAT
ID GTR2_RAT STANDARD: PRT: 522 AA.
AC P12336:
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE GLUCOSE TRANSPORTER TYPE 2, LIVER.
GN SLC2A2 OR GLUT2 OR GLUT-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89003066; PubMed=3048704;
RA Thorens B., Sarkar H.K., Kaback H.R., Lodish H.F.;
RT "Cloning and functional expression in bacteria of a novel glucose
RT transporter present in liver, intestine, kidney, and beta-pancreatic
RT islet cells.";
RL Cell 55:281-290(1988).
CC -!- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY
CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
CC MEMBRANE OF HEPATOCYTES & IS RESPONSIBLE FOR UPTAKE OF GLUCOSE BY
CC THE BETA CELLS; IT MAY COMPRISE PART OF THE GLUCOSE-SENSING
CC MECHANISM OF THE BETA CELL. IT MAY ALSO PARTICIPATE WITH THE
CC CA(+) /GLUCOSE CO-TRANSPORTER IN THE TRANSCELLULAR TRANSPORT OF
CC GLUCOSE IN THE SMALL INTESTINE & KIDNEY.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: PRESENT IN LIVER, INTESTINE, KIDNEY AND
CC BETA-PANCREATIC ISLET CELLS.
CC -!- SIMILARITY: HIGH TO OTHER GLUCOSE TRANSPORTERS.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC
DR EMBL: J03145; AAA41298.1; -
DR PIR: A31556; A31556.
DR InterPro: IPR003663; Sugar transportr.
DR InterPro: IPR003662; sub_transportr.
DR Pfam: PF00083; sugar_tr.1.
DR PRINTS: PR00171; SUGCTRANSPORT.
DR PRINTS: PR00172; GLUCTRANSPORT.
DR PRINTS: PR01191; GLUCTRANSPORT2.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Transport; Glycoprotein;
KW Multigene family.
FT DOMAIN 11 31 POTENTIAL.
FT DOMAIN 32 97 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 98 117 POTENTIAL.
FT TRANSMEM 126 146 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 186 206 POTENTIAL.
FT TRANSMEM 216 236 POTENTIAL.
FT DOMAIN 237 301 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 302 322 POTENTIAL.
FT TRANSMEM 337 357 POTENTIAL.
FT TRANSMEM 368 388 POTENTIAL.
FT TRANSMEM 399 419 POTENTIAL.
FT TRANSMEM 432 452 POTENTIAL.
FT TRANSMEM 460 480 POTENTIAL.
FT DOMAIN 481 522 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 62 62 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 522 AA: 57085 MW: 075AB81E56CF33F7 CRC64:

Query Match 26.6%; Score 135; DB 1; Length 522;
Best Local Similarity 33.0%; Pred No. 1.2e-05;
Matches 32; Conservative 17; Mismatches 38; Indels 10; Gaps 3;
QY 7 LGI-----AMLAVGSMC---LFIAGFAVGMDIPMLNLSSEIFPLHIKGVATGVCLTNM 57
DB 383 LGVLIDKFTWMSYVMTAIFLFVSFEIGDPIPMWVAEFEFSQGPRTALALAFSNM 442
QY 58 FMAFLVTKRNSIMELLRPYGAFWLTAAFCILSVLT 94
DB 443 VCNFIILCFQYIADFLGPY-VFELFAGVVLVFTLT 478

Search completed: February 13, 2002, 22:06:06
Job time: 862 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: February 13, 2002, 22:04:43 ; Search time 136.8 Seconds
(without alignments)
100.509 Million cell updates/sec

Title: US-09-516-493-12
Perfect score: 508
Sequence: 1 EPADVHLGLAWLAVGSMCLF.....RPYGAFWLTAFCILSVLFT 94

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_TREMBL_17:*
2: SP_Archaea:*
3: SP_Bacteria:*
4: SP_Fungi:*
5: SP_Human:*
6: SP_Invertebrate:*
7: SP_Mammal:*
8: SP_MHC:*
9: SP_Organelle:*
10: SP_Phage:*
11: SP_Plant:*
12: SP_Protozoa:*
13: SP_Virus:*
14: SP_Vertebrate:*
15: SP_Unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	508	100.0	478	11	Q9JUZ1
2	508	100.0	479	11	Q9JUA6
3	482	94.9	477	11	Q9JUZ0
4	477	93.9	477	11	Q9JUP4
5	477	93.9	477	11	Q9JUF3
6	457	90.0	477	4	Q9NSC4
7	454	89.4	477	4	Q9NY64
8	222	43.7	433	5	Q9V610
9	220	43.3	507	4	Q9UG03
10	220	43.3	539	5	Q9VU17
11	219	43.1	489	5	Q9V609
12	200.5	39.5	463	10	Q9MAA4
13	189.5	37.3	465	5	Q9V848
14	187.5	36.9	348	10	Q9ZU87
15	186	36.6	444	5	Q9V3G0
16	176.5	34.7	133	10	Q9V6G2
17	171.5	33.8	487	10	Q9FRL3
18	169.5	33.4	621	6	Q9BE72
19	168.5	33.2	510	10	Q65322

20	162.5	32.0	497	5	Q9VNI1	Q9SXB1 arabidopsis
21	162.5	32.0	516	10	Q9LH78	Q9CA87 arabidopsis
22	161.5	31.8	515	10	Q9LH48	Q9M928 arabidopsis
23	160.5	31.6	491	5	Q9VNI2	Q9JLF2 arabidopsis
24	159.5	31.4	514	10	Q39228	Q9S757 arabidopsis
25	159.5	31.4	523	10	Q06312	Q9IM67 arabidopsis
26	158	31.1	493	10	Q9FYG3	Q40373 medicago tr
27	158	31.1	521	10	Q02848	Q91KHI mesembryant
28	157.5	31.0	522	10	Q9SXB1	
29	157	30.9	467	10	Q9CA87	
30	157	30.9	472	10	Q9M928	
31	156.5	30.8	439	10	Q9SCW6	
32	155.5	30.6	490	10	Q9J416	
33	155.5	30.6	518	10	Q9FRT5	
34	155	30.5	457	2	P96742	
35	154.5	30.4	483	10	Q9ZVM0	
36	154	30.3	558	10	Q9FJF2	
37	152.5	30.0	292	10	Q9ST99	
38	152.5	30.0	519	10	Q9Z576	
39	151.5	29.8	433	2	Q9CFH3	
40	151	29.7	582	10	Q23492	
41	150.5	29.6	580	10	Q9C757	
42	150	29.5	734	10	Q9SYQ3	
43	150	29.5	734	10	Q9LM67	
44	149.5	29.4	518	10	Q40373	
45	149	29.3	581	10	Q9LKH1	

ALIGNMENTS

RESULT 1	Q9JUZ1	PRELIMINARY:	PRT:	478 AA.
ID	Q9JUZ1			
AC	Q9JUZ1			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	GLUCOSE TRANSPORTER.			
GN	GLUTX1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.			
NCBI_TaxID	10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-20138191; PubMed-10671487;			
RA	Iberson M.R., Uldry M.A., Thorens B.;			
RT	"Glutx1: A novel mammalian glucose transporter expressed in the			
RT	central nervous system and insulin-sensitive tissues.";			
RL	J. Biol. Chem. 275:4607-4612(2000).			
CC	-I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).			
CC	-I- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.			
DR	EMBL; AJ245935; CAB75729.1; -			
DR	InterPro: IPR003662; sub_transporter.			
DR	InterPro: IPR003663; Sugar_transporter.			
DR	Pfam: PF00083; sugar_tr.1.			
DR	PRINTS; PR00171; SUGRTNSPRT			
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.			
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.			
KW	Transmembrane.			
SEQUENCE	478 AA; 51458 MW; 95841FC1F18C9EE9 CRC64;			
Query Match	100.0%;	Score 508;	DB 11;	Length 478;
Best Local Similarity	100.0%;	Pred. No. 6,4e-42;		
Matches	94;	Conservative 0;	Mismatches 0;	Indels 0;
		Gaps 0;		
OY	1	EPADVHLGLAWLAVGSMCLFAGVNGPILMLSEIFPLHKGATGVCVLTNPMMA 60		
DB	363	EPADVHLGLAWLAVGSMCLFAGVNGPILMLSEIFPLHKGATGVCVLTNPMMA 422		
OY	61	FLVTKFNSIMEILRPYGAFWLTAFCILSVLFT 94		

Db 423 FLVTKFNSIMELRPYGAFWLTAAFCILSVLEFT 456

RESULT 2

Q9JMA6 PRELIMINARY; PRT; 479 AA.

AC Q9JMA6; 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
 DE GLUCOSE TRANSPORTER 8.
 GN GLUT8.

OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;

RN [1]
 RP SEQUENCE FROM N.A.

RA Ishibashi K.;

RT "Molecular cloning of a new putative glucose transporter";
 Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.

DR EMBL; AB03418; BAA94383.1; -

DR InterPro; IPR003663; sub_transporter.

DR Pfam; PF00083; sugar_tr; 1.

DR PRINTS; PR00171; SUGRTNSPORT.

DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.

DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.

KW Transmembrane.
 SQ SEQUENCE 479 AA; 51601 MW; EAB2B678BCDD2FCF CRC64;

Query Match 100.0%; Score 508; DB 11; Length 479;
 Best Local Similarity 100.0%; Pred. No. 6.4e-42;

Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPADVHLGLAWLAVGSMCLFIAGFVAVGMPFIPWLMSEIFPLHKGATGVCVLTNNWMA 60

Db 364 EPADVHLGLAWLAVGSMCLFIAGFVAVGMPFIPWLMSEIFPLHKGATGVCVLTNNWMA 423

QY 61 FLVTKFNSIMELRPYGAFWLTAAFCILSVLEFT 94

Db 424 FLVTKFNSIMELRPYGAFWLTAAFCILSVLEFT 457

RESULT 3

Q9JUZ0 PRELIMINARY; PRT; 477 AA.

AC Q9JUZ0; 01-OCT-2000 (TREMblrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, last sequence update)

DT 01-JUN-2001 (TREMblrel. 17, last annotation update)

DE GLUCOSE TRANSPORTER.

GN GLUTX1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=20138191; PubMed=10671487;

RT Iberson M.R., Uldry M.A., Thorens B.;

RT "GLUTX1: A novel mammalian glucose transporter expressed in the

RT central nervous system and insulin-sensitive tissues.";

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.

DR EMBL; AJ245936; CAB75719.1; -

DR MGD; MGI:1860103; GLUTX1.

DR InterPro; IPR003662; sub_transporter.

DR InterPro; IPR003663; Sugar_transporter.

DR Pfam; PF00083; sugar_tr; 1.

DR PRINTS; PR00171; SUGRTNSPORT.

DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.

DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.

KW Transmembrane.

SQ SEQUENCE 477 AA; 51578 MW; 14FF23DB9060DAF1 CRC64;

Query Match 94.9%; Score 482; DB 11; Length 477;
 Best Local Similarity 91.5%; Pred. No. 2.2e-39;

Matches 86; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 EPADVHLGLAWLAVGSMCLFIAGFVAVGMPFIPWLMSEIFPLHKGATGVCVLTNNWMA 60

Db 362 EPADVHLGLAWLAVGSMCLFIAGFVAVGMPFIPWLMSEIFPLHKGATGVCVLTNNWMA 421

QY 61 FLVTKFNSIMELRPYGAFWLTAAFCILSVLEFT 94

Db 422 FLVTKFNSIMELRPYGAFWLTAAFCILSVLEFT 455

RESULT 4

Q9JUP4 PRELIMINARY; PRT; 477 AA.

AC Q9JUP4; 01-OCT-2000 (TREMblrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, last sequence update)

DT 01-JUN-2001 (TREMblrel. 17, last annotation update)

DE GLUCOSE TRANSPORTER 8.

GN GLUTX1 OR GLUT8.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA TISSUE=TESTIS;

RT MEDLINE=20283667; PubMed=10821868;

RT "Doege H., Scherrenn A., Bahrenberg C., Brauers A., Joost H.G.;

RT "GLUT8: A novel member of the sugar transport facilitator family with

RT glucose transport activity.";

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.

DR EMBL; Y17802; CAB89815.1; -

DR MGD; MGI:1860103; GLUTX1.

DR InterPro; IPR003662; sub_transporter.

DR InterPro; IPR003663; Sugar_transporter.

DR Pfam; PF00083; sugar_tr; 1.

DR PRINTS; PR00171; SUGRTNSPORT.

DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.

DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.

KW Transmembrane.

SQ SEQUENCE 477 AA; 51507 MW; 59A985B52FADF478 CRC64;

Query Match 93.9%; Score 477; DB 11; Length 477;
 Best Local Similarity 90.4%; Pred. No. 6.7e-39;

Matches 85; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 EPADVHLGLAWLAVGSMCLFIAGFVAVGMPFIPWLMSEIFPLHKGATGVCVLTNNWMA 60

Db 362 EPADVHLGLAWLAVGSMCLFIAGFVAVGMPFIPWLMSEIFPLHKGATGVCVLTNNWMA 421

QY 61 FLVTKFNSIMELRPYGAFWLTAAFCILSVLEFT 94

Db 422 FLVTKFNSIMELRPYGAFWLTAAFCILSVLEFT 455

RESULT 5

Q9JIF3 PRELIMINARY; PRT; 477 AA.


```

AC 09JF3:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE GLUCOSE TRANSPORTER GLUT8.
GN GLUTX1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RA MEDLINE=20319023; Pubmed=10860996;
RA Carayannopoulos M.O., Chi M.M.Y., Cui Y., Pingsterhaus J.M.,
RA McKnight R.A., Mueckler M., Devaskar S.U., Moley K.H.;
RT "GLUT8 is a glucose transporter responsible for insulin-stimulated
RT glucose uptake in the blastocyst.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:7313-7318(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC EMBL: AF232061; AAF78366.1; -.
DR MGD: MGI:1860103; GLUTX1.
DR InterPro: IPR003662; sub_transporter.
DR InterPro: IPR003663; Sugar_transporter.
DR Pfam: PF00083; sugar_tr.1.
DR PRINTS: PR00171; SUGRTNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane.
SQ SEQUENCE 477 AA; 51523 MW; A3753FB34E452F9A CRC64;

Query Match          93.9%; Score 477; DB 11; Length 477;
Best Local Similarity 90.4%; Pred. No. 6,7e-39;
Matches 85; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 1 EPADVHLGLAVLAVGSMCLFIAGFAVGWGPFPWLLMSEIFPLHIKVGATGCVLTWFWMA 60
DB 362 QPVDASVGLAVLAVGSMCLFIAGFAVGWGPFPWLLMSEIFPLHIKVGATGCVLTWFWMA 421
OY 61 FLVTKFNSIMELIRPYGAFWLTAFCLISVLT 94
DB 422 FLVTKFSSIMELIRPYGAFWLTAFCLISVLT 455

RESULT 6
O9NSC4 PRELIMINARY: PRT: 477 AA.
AC O9NSC4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE GLUCOSE TRANSPORTER 8.
GN GLUT8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA MEDLINE=20283667; Pubmed=10821868;
RA Doerge H., Schuerman A., Bahrenberg C., Brauers A., Joost H.G.;
RT "GLUT8: A novel member of the sugar transport facilitator family with
RT glucose transport activity.";
RL J. Biol. Chem. 275:16275-16280(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC EMBL: Y17801; CAB89809.1; -.
DR InterPro: IPR003662; sub_transporter.
DR InterPro: IPR003663; Sugar_transporter.
DR Pfam: PF00083; sugar_tr.1.

```

```

DR PRINTS: PR00171; SUGRTNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane.
SQ SEQUENCE 477 AA; 50792 MW; 0B480F94B40AE76 CRC64;

Query Match          90.0%; Score 457; DB 4; Length 477;
Best Local Similarity 85.1%; Pred. No. 6e-37;
Matches 80; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

OY 1 EPADVHLGLAVLAVGSMCLFIAGFAVGWGPFPWLLMSEIFPLHIKVGATGCVLTWFWMA 60
DB 362 QPVDASVGLAVLAVGSMCLFIAGFAVGWGPFPWLLMSEIFPLHIKVGATGCVLTWFWMA 421
OY 61 FLVTKFNSIMELIRPYGAFWLTAFCLISVLT 94
DB 422 FLVTKFSSIMELIRPYGAFWLTAFCLISVLT 455

RESULT 7
O9NY64 PRELIMINARY: PRT: 477 AA.
AC O9NY64;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE GLUCOSE TRANSPORTER.
GN GLUTX1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20138191; Pubmed=10671487;
RA Ibberson M.R., Uldry M.A., Thorens B.;
RT "GLUTX1: A novel mammalian glucose transporter expressed in the
RT central nervous system and insulin-sensitive tissues.";
RL J. Biol. Chem. 275:4607-4612(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
CC EMBL: AJ245937; CAB75702.1; -.
DR InterPro: IPR003662; sub_transporter.
DR InterPro: IPR003663; Sugar_transporter.
DR Pfam: PF00083; sugar_tr.1.
DR PRINTS: PR00171; SUGRTNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane.
SQ SEQUENCE 477 AA; 50858 MW; 39DA9E2DBFF8D3E2 CRC64;

Query Match          89.4%; Score 454; DB 4; Length 477;
Best Local Similarity 84.0%; Pred. No. 1.2e-36;
Matches 79; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

OY 1 EPADVHLGLAVLAVGSMCLFIAGFAVGWGPFPWLLMSEIFPLHIKVGATGCVLTWFWMA 60
DB 362 QPVDASVGLAVLAVGSMCLFIAGFAVGWGPFPWLLMSEIFPLHIKVGATGCVLTWFWMA 421
OY 61 FLVTKFNSIMELIRPYGAFWLTAFCLISVLT 94
DB 422 FLVTKFSSIMELIRPYGAFWLTAFCLISVLT 455

RESULT 8
O9V610 PRELIMINARY: PRT: 433 AA.
AC O9V610;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)

```


SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abill J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dudin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pitman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shee B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RL "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -i- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -i- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL: AE003540; AAF49874.1; -
 DR Flybase: FBgn0036316; CG10960.
 DR InterPro: IPR003662; sub_transporter.
 DR InterPro: IPR003663; Sugar_transporter.
 DR Pfam: PF00083; sugar_trt.1
 DR PRINTS: PR00171; SUGARTRANSPORT.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane.
 SQ SEQUENCE 539 AA; 58413 MW; A2F4814DBF676BBD CRC64;

Query Match 43.3%; Score 220; DB 5; Length 539;
 Best Local Similarity 44.8%; Pred. No. 9,1e-14;
 Matches 39; Conservative 18; Mismatches 30; Indels 0; Gaps 0;

7 LGTGLAVGSMCLTAGAVGCGPIPLMLMSEIPPLHIKGVATGCVLTNNMFAPLVYTK 66
 DB 419 VSLGMLPVAASLCGLTFLIMFSIGGPPVPLMGMELFATDIKFGAGSLAGTSNMLLAFLVYTKT 478

67 FNSIMEILRPYGAFWLTAFCILSYLF 93
 DB 479 FVNLDNGIGIGGTFWLFGLTAVGVIF 505

RESULT 11
 09V609 PRELIMINARY: PRT: 489 AA.
 AC 09V609:
 DT 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)

01-JUN-2001 (TRENBLREL. 17, Last annotation update)
 DE CG7801 PROTEIN.
 GN CG7801.
 OS *Drosophila melanogaster* (Fruit Fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abill J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dudin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pitman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shee B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RL "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -i- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -i- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL: AE003825; AAF58631.1; -
 DR Flybase: FBgn003663; CG7801.
 DR InterPro: IPR003662; sub_transporter.
 DR InterPro: IPR003663; Sugar_transporter.
 DR Pfam: PF00083; sugar_trt.1
 DR PRINTS: PR00171; SUGARTRANSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 2.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane.
 SQ SEQUENCE 489 AA; 53026 MW; D26436DAD69723C6 CRC64;

Query Match 43.1%; Score 219; DB 5; Length 489;
 Best Local Similarity 44.3%; Pred. No. 1e-13;
 Matches 39; Conservative 14; Mismatches 33; Indels 2; Gaps 1;

6 HUGTGLAVGSMCLTAGAVGCGPIPLMLMSEIPPLHIKGVATGCVLTNNMFAPLVYTK 65
 DB 364 HLG-WLPLTFCVYIYILGSLGCGPIPLMLMGMELTLPKIRGSAASAVATFNNMCTFVYTK 421

66 EFNSIMEILRPYGAFWLTAFCILSYLF 93

```

Db 422 TFODLTVMAGGAFWLTGALCEVGLTF 449
RESULT 12
ID Q9MAA4 PRELIMINARY; PRT: 463 AA.
AC Q9MAA4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE SUGAR TRANSPORTER.
GN T12H1.11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Cressy T.H., Haas B., Wu D.,
RA Roming C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC T12H1 genomic sequence.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL: AC009177; AAF27021.1; -
DR InterPro: IPR003662; sub_transporter.
DR InterPro: IPR003663; sugar_transporter.
DR Pfam: PF00083; sugar_tr; 1.
DR PRINTS: PR00171; SUGSTRNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW Sugar transporter; Transmembrane.
SQ SEQUENCE 463 AA; 50148 MW; 82B93947AE79DAC0 CRC64;

Query Match 39.5%; Score 200.5; DB 10; Length 463;
Best Local Similarity 41.5%; Pred. No. 6.3e-12;
Matches 34; Conservative 23; Mismatches 24; Indels 1; Gaps 1;

QY 12 LAVSGMCLFIAGFAVGMCPITWLMSEIFPLIKGVANGCVLTNMFMAFLTKFENSIM 71
DB 362 LAVSGVLVYISGFSIGMGALPVMVISEIFPLINKGAGLAVTVVMVLSLVSTFENFLM 421
QY 72 ELIRPYGAFWLTAAFCILSVLF 93
DB 422 -TWSPHGTFYVYGVCVLAITF 442

RESULT 13
ID Q9V848 PRELIMINARY; PRT: 465 AA.
AC Q9V848;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CG6484 PROTEIN.
GN CG6484.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang O., Chen L.X.,

```

```

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegun C.,
RA Jajali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy C., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL: AE003802; AAF57829.1; -
DR FlyBase: FBgn0034247; CG6484.
DR InterPro: IPR003662; sub_transporter.
DR InterPro: IPR003663; sugar_transporter.
DR Pfam: PF00083; sugar_tr; 1.
DR PRINTS: PR00171; SUGSTRNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane.
SQ SEQUENCE 465 AA; 50271 MW; 004B4CACF406BDB CRC64;

Query Match 37.3%; Score 189.5; DB 5; Length 465;
Best Local Similarity 34.0%; Pred. No. 7.5e-11;
Matches 32; Conservative 24; Mismatches 37; Indels 1; Gaps 1;

QY 1 EPADVHLGLAWLAVGSMCLFIAGFAVGMCPITWLMSEIFPLIKGVANGCVLTNMFMA 60
DB 340 DPASMD-NEGWLPISISICFIIFFSIGFGVPLVALELSDVSAAGIAGTSMWSA 398
QY 61 FLVTKFENSIMEILRPYGAFLTAFCILSVFT 94
DB 399 FVTLTFLTKSSIGPGTFWFTTAIVIAFYYS 432

RESULT 14
ID Q9Z087 PRELIMINARY; PRT: 348 AA.
AC Q9Z087;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE SUGAR TRANSPORTER.
GN T9J23.17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

```

OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 RX NCBI_TaxID=3702;
 RA [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Keul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
 RA Barnstead M.E., Mason T.M., Bowman C.L., Romling C.M., Benito M.,
 RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
 RA Fraser C.M., Venter J.C.;
 RA "Arabidopsis thaliana chromosome II BAC T9J23 genomic sequence."
 RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL: AC006072; AAD13706.1; -;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr.1.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
 KW Sugar transport; Transmembrane
 SQ SEQUENCE 348 AA: 36899 MW: C9A7C75133FF69BF CRC64:

Query Match 36.9%; Score 187.5; DB 10; Length 348;
 Best Local Similarity 43.2%; Pred. No. 8.9e-11;
 Matches 38; Conservative 18; Mismatches 31; Indels 1; Gaps 1;

OY 6 HGLAMLVGSMCLFIAGFAVGWGPFIPLMLMSEIFPLHIKGVATGCVLTNNFMAFLVTKR 65
 DB 240 HEAVPLVAVGIMVYIGSFSGAGMGPVYVSEIFPINIKGVAGGMATLVNFGMAVSY 299
 OY 66 EFNSIMELIRPYGAFWLTAFNCILSVLF 93
 DB 300 TFNPLMS-WSSYGFPLIYAAMALAIWF 326

RESULT 15
 ID 09V3G0 PRELIMINARY; PRT: 444 AA.

AC 09V3G0;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE SUT4 PROTEIN.
 GN SUT4 OR CG1380.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NX NCBI_TaxID=7227;
 RA [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaratilake P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J.J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Sviderskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodard T., Wortley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Andersson Escher S., Lohlahti M., Rasmuson-Leestander A.;
 RT "Sugar transporters in Drosophila melanogaster."
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL: AE003831; AAF58847.1; -;
 DR EMBL: AF199486; AAF13274.1; -;
 DR Flybase: FBgn0028560; sut4.
 DR InterPro: IPR003662; sub_transporter.
 DR InterPro: IPR003663; Sugar_transporter.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGARTRANSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Sugar transport; Transmembrane
 SQ SEQUENCE 444 AA: 48028 MW: D18068459A9861A CRC64:

Query Match 36.6%; Score 186; DB 5; Length 444;
 Best Local Similarity 37.2%; Pred. No. 1.6e-10;
 Matches 32; Conservative 23; Mismatches 31; Indels 0; Gaps 0;

OY 8 GLAMLVGSMCLFIAGFAVGWGPFIPLMLMSEIFPLHIKGVATGCVLTNNFMAFLVTKR 67
 DB 334 GVGWLTALVAVCFIIGFSGFPGVPMMLMAELFADVVKALAGSIAGTWMCFATITLTF 393
 OY 68 NSIMELIRPYGAFWLTAFNCILSVLF 93
 DB 394 PVLNDITGATGCFATPGFAVAFWF 419

Search completed: February 13, 2002, 22:04:44
 Job time: 865 sec

Thu Feb 14 07:44:31 2002

us-09-516-493-12.rspt

Page 8


```

465 GGGCCCCGCTACATCTCCGAATCGCCTACCCAGACAGTCCGGGGTTCG 514
134 euglyseCysValGlnleuMetValValIglYlleLeuLeuAlaLyr 150
|||||
515 TCGGCTCTCTGTGACAGCTAAATGCTGCTGCGGATCTCTCTGCGCTAC 564
151 leuAlaGlyTPValLeuGluTPArgTPheuaIaValLeuGlyCysVa 167
|||||
565 CTGGCAGGCTGGGTGAGTAGTGGCGCTGGCTGTGCTGGCTGGCTGGT 614
167 lProGSerleuMetleuLeuMetCysPheMetProGluThrProA 184
|||||
615 GCCCCCTCCCTCATTGCTGCTTCATGCTGCTTATGCCGAGACCCGCG 664
184 rphLeuLeuThrGlnHisArgArgGlnGlnAlaIle...AlaLeuArg 199
|||||
665 GCTTCCGCTGACACACAGGCGCCAGAGGCGCATGGCCCTCGG 714
200 PheLeuTPolyserylGlnGlyTPArgLuAspProProlleGlyAlaG 216
|||||
715 TTCTCTGGGGGCTCCGAGAGGCGTGGGAAGACCCCCCATCGGGGCTGA 764
216 uGlnSerPheHisleuAlaLeuLeuArgGlnProGlyIleTyrlsProp 233
|||||
765 GCAGAGCTTTCACCTGCGCTGCTGGCGGAGCCCGCATCTACAAAGCCT 814
233 heileileGlyValSerleuMetAlaPheGlnGlnleuSerGlyValAsn 249
|||||
815 TCATCATCGGCGCTCCCTGATGCGCTTCACGACGCTGTGGGGGTAC 864
250 AlaValMetPheTyrlAlaGluThrIlePheGluAlaAlaLysPheLys 266
|||||
865 GCGGTATGTTCTATGACAGACCATTTGAAGAGGCCAAGTTCAAGA 914
266 pSerSerleuAlaSerValValIglYlleGlnValleuPheThr 283
|||||
915 CAGCAGCTGCGCTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 964
283 lAValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuVal 299
|||||
965 CTGTGCGGCTCTCATCATGACAGAGGCGGAGGCGCTGCTGCTGCTG 1014
300 leuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaLyrPh 316
|||||
1015 TTGTACAGGTGTGTCATGTTGACACAGAGTCCCTGCGGCTTACTT 1064
316 eLysleuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerA 333
|||||
1065 CAAGCTGACCCAGGGTGGCCCTGGGCACTCTCGCACGTGGCCATCTCG 1114
333 lAProValSerAlaGlnProValAspAlaSerValGlyLeuAlaThrLeu 349
|||||
1115 GCGCTGTCTCTGACAGCTCTTGTATGCCAGCGTGGGCTGCGCTGCTG 1164
350 AlaValIglYSerMetCysLeuPheIleAlaGlyPheAlaValAlaGlyTP 366
|||||
1165 GCGGTGGGCGAGATGCTGCTTTCATCGCGGCTTGGCGTGGGCTGGG 1214
366 yProIleProTPleuLeuMetSerGluIlePheProLeuHisValLysG 383
|||||
1215 GCCATCCCTCGGCTGCTCATGTAGAGATCTTCCCTCGCATGTCAGAG 1264
383 lYValAlaThrGlyIleCysValLeuThrAsnTPleuMetAlaPheLeu 399
|||||
1265 GCGTGGCGACAGGATCTGCGCTCCACCAACTGGCTCATGGCTTCTCT 1314
400 ValThrIysGluPheSerSerleuMetGluValleuArgProTyrlGlyAl 416
|||||
1315 GTAGACAGAGGATTCAGACGCTCTGAGAGTCTCTCAGGCGCTTAGGAG 1364
416 aPheTPleuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuP 433
|||||
1365 CTCTGCGCTGCTCCGCTTCTGTCATCTTCAGTGTCTTCTTCACTTGT 1414

```

```

433 heCysValProGluThrLysGlyLysThrLeuGlnGlnIleThrAlaHis 449
|||||
1415 TCTGTGTCCCTGGAATAAGAAAGACTGTGAAACAATACAGGCCCAT 1464
450 PheGluGlyArg 453
|||||
1465 TTTCAGGGGCGGA 1476

seq_name: gb-pr: HSA245937
seq_documentation_block:
LOCUS HSA245937 1873 bp mRNA PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA for glucose transporter (GLUTX1 gene).
ACCESSION AJ245937
VERSION AJ245937.1 GI:7018305
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1873)
AUTHORS Ibberson,M., Uldry,M. and Thorens,B.
TITLE GLUTX1, a novel mammalian glucose transporter expressed in the
central nervous system and insulin-sensitive tissues
JOURNAL J. Biol. Chem. 275 (7), 4607-4612 (2000)
MEDLINE 20138191
REFERENCE 2 (bases 1 to 1873)
AUTHORS Ibberson,M.R.
TITLE Direct Submission
SUBMITTED (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
CH-1005, SWITZERLAND
FEATURES
source
location/Dualifiers
1..1873
/organism="Homo sapiens"
/db_xref="taxon:9606"
4..1437
/gene="GLUTX1"
4..1437
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CA875702.1"
/db_xref="GI:7018306"
/translation="MTPEDEETQPLIGPPGSGAPRGRRVFLAFAAALGSLSPGAL
GYSPPALPSLQRAAPAPRLDAAASWFGAVVTLGAAGVGLGWLVDPRGRKLSLL
CSYFVAGFVITVIAADQVMMILGRLITGLACVASLVAPEYTSIETAYPVRGLISC
VOLMNVVGIILAVLAGVLEWRMLAVIGCVPSIMLLMGFMETPRELLTHORROEA
MAALRFLEMGSEOGMEDPPIGAEOSPHALLROPGIYKPTIIGVSLAFOQLSYNAVY
AFVAFETFEERAKFRDSSLASAVYGVIVLTFVAAALIMDRAGRLLVLVSGVNVFST
AFGAFYKRLTGGGNSHVAISAPVSAQPDVASVGLAWLAVGMKCLFIAGFAVAGMPI
PMLMSIEPLHAKGVATGICVLTNNIMLAVLTKRESSIMELVLRPYGAWLASAPCIF
SVLTLECEVEIKLEQITAFHEGR"
BASE COUNT 279 a 621 c 573 g 400 t
ORIGIN

alignment_scores:
Quality: 2301.50 Length: 454
Ratio: 5.092 Gaps: 1
Percent Similarity: 99.559 Percent Identity: 99.119

alignment_block:
US-09-516-493-7 x HSA245937 ..
Align seg 1/1 to: HSA245937 from: 1 to: 1873
1 ArgATGValPheLeuAlaAlaPheAlaAlaIleGlyProLeuSerPh 17
|||||
73 CGCGCGCTCTCTGCGCGCTTGGCCGCTGCGCTGGGCGCCACTCAGCTT 122
|||||
17 eGlyPheAlaLeuGlyTyrlSerSerProAlaIleProSerleuGlnArg 34

```


alignment_scores:

Quality: 2301.50 Length: 454
 Ratio: 5.092 Gaps: 1
 Percent Similarity: 99.559 Percent Identity: 99.119

alignment_block:

US-09-516-493-7 x AX076667 ..

Align seg 1/1 to: AX076667 from: 1 to: 2217

```

1 ArgArgValPheLeuAlaIaIaPheAlaIaIaLeuGlyProLeuSerPh 17
  |||
417 CGCGCGCTTCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCT 466
17 eGlyPheAlaLeuGlyTyrSerSerProAlaIaIaProSerLeuGlnAra 34
  |||
467 CGGCTTGGCGCTCGCTACAGACTCCCGGCTATCTTACCTGACGAGCG 516
34 AlaIaProProAlaIaProArgLeuAspAspAlaIaIaSerTrpPheGly 50
  |||
517 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCGCTGTCGG 566
51 AlaValValThrLeuGlyAlaIaIaIaGlyValLeuGlyTyrPhe 67
  |||
567 GCTGTCTGACCTCGGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCTGCT 616
67 uValAspArgAlaGlyArgLeuSerLeuLeuLeuLeuLeuLeuLeuLeu 84
  |||
617 GGTGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCT 666
84 heValAlaGlyPheAlaValIleThrAlaIaIaGlnAspValTrpMetLeu 100
  |||
667 TCGTGGCGCGCTTGGCTGCTATCATCCGCGCGCGCGCGCGCGCGCTG 716
101 LeuGlyValArgLeuLeuThrGlyLeuAlaCysGlyValAlaIaSerLeu 117
  |||
717 CTGGGGGGCGCGCTCGCTACCGCGCGCGCTGCGCTGCTGCTGCTGCTG 766
117 AlaProValTyrIleSerGluIleAlaIaIaIaIaIaIaIaIaIaIa 134
  |||
767 GGTGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCT 816
134 euGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaIa 150
  |||
817 TCGGCTCGCTGCTGCGAGCTAATGCTGCTGCTGCTGCTGCTGCTGCTG 866
151 LeuAlaGlyTrpValLeuGluTrpArgTrpLeuAlaValLeuGlyCys 167
  |||
867 CTGGAGCGCTGGGCTGCTGAGTGGCGCTGCTGCTGCTGCTGCTGCTG 916
167 IProProSerLeuMetLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 184
  |||
917 GCGCGCGCTCGCTATGCTCTTCTATGCTCTTCTATGCTCTTCTATGCT 966
184 rPheLeuLeuThrGlnIleSerArgArgGlnGluAlaIle..AlaLeuArg 199
  |||
967 GCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1016
200 PheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAla 216
  |||
1017 TTCCGTGTGGGCTCGAGAGCGGCTGGGAAGACCCCGCATCGGGGCTGA 1066
216 uGlnSerPheHisLeuAlaIaLeuLeuArgGlnProGlyIleTyrLeuPro 233
  |||
1067 GCAGAGCTTTTACCTGCGCTGCTGCTGCGGAGCGCGCATCTACAGCCCT 1116
233 heIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsn 249
  |||
1117 TCATCATCGGCGCTTCCCTGATGGCTTCCACACACTGTCGGGGGCTAAC 1166
250 AlaValMetPheTyrAlaGluThrIlePheGluGluAlaIaIaIaIaIa 266
  |||
1167 GCCGTCACTGTTCTATGAGAGACCATCTTTGAAAGAGCCCAAGTTCAGGA 1216

```

```

266 pSerSerLeuAlaSerValValValGlyValIleGlnValLeuPheThrA 283
  |||
1217 CAGCAGCTGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1266
283 LaValAlaIaIaLeuIleMetAspArgAlaGlyArgArgLeuLeuVal 299
  |||
1267 CTGTGGCGGCTCTCATCTGAGACAGAGCGGCGGAGGCTGCTGCTGCTG 1316
300 LeuSerGlyValValMetValPheSerThrSerIaPheGlyAlaTyrPh 316
  |||
1317 TTGTGAGGTGTGATCATGATGATGATGATGATGATGATGATGATGATG 1366
316 eLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerA 333
  |||
1367 CAAGCTTACCCAGAGGCTGCGCTGCGCAACTCTCGCACGCTGGCCATCT 1416
333 IaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaIaTrpLeu 349
  |||
1417 CGCTGTCTGTGACAGACTGTGATGATGATGATGATGATGATGATGATG 1466
350 AlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpG 366
  |||
1467 GCGGTGGCAACATGTCCTCTTCATGCGCGCTTGGCGGTGGGCTGGGG 1516
366 YProIleProTrpLeuLeuMetSerGluIlePheProLeuHisValIysG 383
  |||
1517 GCCATCCCTGCTGCTCTCATGTCAGAGATCTTCTGCTGCTGCTGCTG 1566
383 LyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeu 399
  |||
1567 GCGTGGGAGACGACATCTGCTCTCACCAACTGGCTCATAGGCTTTTCT 1616
400 ValThrLysGluPheSerSerLeuMetGluValLeuArgProTrpGlyAl 416
  |||
1617 GTAGCCAGAGATTGACAGACCCATGAGAGTCTCTGACGCCCTATGAGAG 1666
416 aPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeu 433
  |||
1667 CTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1716
433 heCysValProGluTrpLysGlyLysThrLeuGlnIleThrAlaHis 449
  |||
1717 TCTGTCTCCCTGAATTAAGGAAAGACTGTGGAACAAATCACAGCCAT 1766
450 PheGluGlyArg 453
  |||
1767 TTGAGGGGCGA 1778

```

seq_name: gb_pr:HSAl7801

seq_documentation_block: 1508 bp mRNA PRI 13-MAY-2000

LOCUS HSAl7801

DEFINITION Homo sapiens mRNA for glucose transporter 8 (GLUT8 gene).

ACCESSION Y17801

VERSION Y17801.1 GI:7688145

KEYWORDS glucose transporter 8; GLUT8 gene.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1508)

AUTHORS Doege,H., Schumann,A., Bahrenberg,G., Brauers,A. and Joost,H.G.

TITLE GLUT8, a novel member of the sugar transport facilitator family with glucose transport activity

JOURNAL J. Biol. Chem. 275 (21), 16275-16280 (2000)

MEDLINE 20283667

REFERENCE 2 (bases 1 to 1508)

AUTHORS Joost,H.G.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-1998) H. Joost, Institute of Pharmacology and Toxicology, Technical University Aachen, Wendlingweg 2, D-52057 Aachen, FRG

1446 TTTGAGGGCGA 1457

seq_name: gb_pat:AX076669

seq_documentation_block:

LOCUS AX076669 2087 bp DNA PAT 06-FEB-2001
 DEFINITION Sequence 3 from Patent WO0104145.
 ACCESSION AX076669
 VERSION AX076669.1 GI:12711200
 KEYWORDS
 SOURCE Rattus sp.
 ORGANISM Rattus sp.
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 2087)

AUTHORS Thorens,B., Ibberson,M. and Uldry,M.
 TITLE Gluc polypeptide family and nucleic acids encoding same
 JOURNAL Patent: WO 0104145-A 3 18-JAN-2001;
 University of Lausanne (CH)

FEATURES Location/Qualifiers

source

1..2087
 /organism="Rattus sp."
 /db_xref="taxon:10118"
 30..1466
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC28496.1"
 /db_xref="GI:12711201"

CDS

translation="MSPEPQETQPLNSPGRADGRREVLATFAALGPIISFGFAL
 CYSSPAIRSLRTPAPLRLGDTASWFGAVVTLGAAGVIGWLLDRAGKLSLL
 CTVPVGTAVTAAVDVWMLGRLTLGACVLAAGVISEIAYPAVGKILSGC
 VOLMVTGILAVAVWLEWMLAVLGCPTLMLLMCTWPEPRFLITQHOGEA
 MVALPFLWSEGEPEPVGAHOGFOLAMLRPVHPLIIGICLAMEFOOCSNAT
 MEYANTIPEEAKFEDKSSLASVYGIQVLTFAVALIMDRAGKILLASGVINFSM
 SAFGTFKLTGSPSSSHVGLVPIISAPADYHGLGLMVAAGSMCLTAGAVGWGP
 IPLLMSRILPRLHKVATGCVLTNNMFMAFLVTKFNSIMEILLRPGAFMLTAFCI
 LSVFTLTFPEPKRLEQLTAHFEGR"
 BASE COUNT 355 a 673 c 580 g 479 t
 ORIGIN

alignment_scores:

Quality: 2057.00 Length: 455
 Ratio: 4.633 Gaps: 2
 Percent Similarity: 97.582 Percent Identity: 85.714

alignment_block:

US-09-516-493-7 x AX076669

Align seg 1/1 to: AX076669 from: 1 to: 2087

1 ArgAtgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPh 17
 |||||
 99 CGCGGGCTCTTCTTCACCTCGCGCGCGCTGGCGCTGACGCTT 148
 17 egiyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuAlaArga 34
 |||||
 149 CGGCTTGGCGCTTGCTACAGCTCCCGCCCATCCCGACCTGCGCGCGCA 198
 34 lAlaLProProAlaProArGLeuAspAlaAlaAlaSerTrpPhegly 50
 |||||
 199 CGCGACCCCTCCCTACCGCTCGAGACAGCTGCGGCTCTGCTGCGG 248
 51 AlavaIValThrLeuGlyAlaAlaAlaGlyValLeuGlyGlyTyrPle 67
 |||||
 249 GCCGTCGTGACCTCGCGCTCGGAGAGGCGTGGCTGGCGGCTGCT 298
 67 uValAspArgAlaGlyArgGlySerLeuSerLeuLeuGlySerValProp 84
 |||||
 299 CTTGACCGCTGACGGGCGCAAGCTGAGCTCTCTGTCACACCGGCT 348
 84 heValAlaGlyPheAlaValAlaIleThrAlaAlaGlyAspValTrpMetLeu 100

|||||
 349 TCGTACCGGTTTGGTGTATCATACCGCGCGCGGAGATGTGTGATGCTG 398
 101 LeuGlyGlyArgLeuLeuThrGlyLeuAlaGlySerGlyValAlaSerLeuVa 117
 |||||
 399 CTCGAGGCGCGCTCTCTACCGGCTACCGGCTGCGAGCGCTCATAGT 448
 117 lAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeu 134
 |||||
 449 GGCACCGGCTATATCTGGAATCGCCTACCGAGCGCTTGAGAGACTGC 498
 134 euGlySerGlyValGlnLeuMetValValGlyIleLeuLeuAlaTyr 150
 |||||
 499 TCGGCTCTGTGTGACAGCTGATGGTTGTCACTGGACATCTTGGCTAC 548
 151 LeuAlaGlyTrpValLeuGluTrpArgTrpLeuAlaValLeuGlyGlyVa 167
 |||||
 549 GTGGCAGGCTGGCTCTAGAAATGGCGTGGCTGTGCTGGGCTGTGT 598
 167 lProProSerLeuMetLeuLeuLeuMetCysPheMetProGluThrPro 184
 |||||
 599 GCCCCCGACCCCTCATGCTGCTGTCATGTCTCATATGCGCGAGACCC 648
 184 rPheLeuLeuThrGlnHisArgArgGlnAlaAlaIle...AlaLeuArg 199
 |||||
 649 GCTTCTCTCTGACGCAACACAGTACAGAGGAGCCATGCTGCTGGCG 698
 200 PheLeuTrpGlySerGlnGlnGlyTrpGluAspProProIleGlyAla 216
 |||||
 699 TTCTTGAGGCGCTCTGAGAGGCTGGAAGAGCCCGCTTGTGGGCTGA 748
 216 u...GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLys 232
 |||||
 749 GACACAGGCTTCCACTGCTGCTATGCTGAGCGCCCTGCTGTCACAGC 798
 232 roPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyVal 248
 |||||
 799 CCTCATCATCGGCATTTGCTCATGCTCTTCAGAGAGCTGTACAGGGG 848
 249 AsnAlaValMetPheTyrAlaGluThrIlePheGlnGluAlaLysPhe 265
 |||||
 849 AACCATCATGTTCTATGCAACACCATTTGAGAGGCCCAAGTTCAA 898
 265 sAspSerSerLeuAlaSerValValValGlyValIleGlnValLeuPhe 282
 |||||
 899 GAGACAGAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 948
 282 hrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeu 298
 |||||
 949 CTGCTGTGCGGCGCTCATCATGAGACAGAGGCGGAACCTGCTTCTG 998
 299 ValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAla 315
 |||||
 999 GCTTGTGCGGTGTATCATGTTTTCAGCATGACCGCTTGTGTACCTA 1048
 315 rPheLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIle 332
 |||||
 1049 CTTCAAACTGACCCAGAGAGGCGCCAGCACTCTCTCCATGTAGGCTTC 1098
 1099 TGTGCTCCCATCTCCAGAGAGCTGCTGATGTTACCTGCGGCTGCGCT 1148
 332 erAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrp 348
 |||||
 1149 TGGCTGTAGAGCATGTGCTTTCATCGCTGTTTTCAGTAGAGCTG 1198
 349 LeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTr 365
 |||||
 365 pGlyProIleProTrpLeuLeuMetSerGluIlePheProLeuHisVal 382
 |||||
 1199 GGGACCATTCCTGCTGCTCATGTCTGAGATCTTCCCTCTGACATCA 1248
 382 ysgIyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPhe 398
 |||||

```
1249 AGGCTGTGCTACCGCGCTGTCTCTCACCACCTGTTTCATGGCCTTT 1298
399 LeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlu 415
|||||.....
1299 CTGGTGACCAAGAGTTTAACGATCATGAGATGATCCTCAGACCCCTACGG 1348
415 ValAlaThrPheLeuAlaSerAlaPheCysIlePheSerValIlePheThrL 432
|||||.....
1349 CGCCTTCTGGCTCACCGCTCTGTATCCTCAGCGCTTTTTCACGC 1398
442 eupPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThrAla 448
|||||.....
1399 TCACCTTTGTCCTCGAGACTTAAGCAGACTGTGGAACAATCACAGCC 1448
449 HisPheGluGlyArg 453
|||||.....
1449 CATTCGAGGAGCAG 1463

seq_name: gb_ro:RNO245935
seq_documentation_block:
LOCUS RNO245935 2087 bp mRNA 18-FEB-2000
DEFINITION Rattus norvegicus mRNA for glucose transporter (GLUTX1 gene).
ACCESSION AJ245935
VERSION AJ245935.1 GI:7018604
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2087)
AUTHORS Ibberson,M., Uldry,M. and Thorens,B.
TITLE GLUTX1, a novel mammalian glucose transporter expressed in the
JOURNAL J. Biol. Chem. 275 (7), 4607-4612 (2000)
MEDLINE 2 (bases 1 to 2087)
REFERENCE Ibberson,M.R.
AUTHORS Direct Submission
TITLE Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
JOURNAL and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
CH-1005, Switzerland
FEATURES
Source
1..2087
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
30..1466
/gene="GLUTX1"
30..1466
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CAB5729.1"
/db_xref="GI:7018605"
/translacion="MSPEDPQETQPLRSPGARPGRRVFLTAALGPISTFGL
GYSPPALSRTPALRLGDTAAWFGAVVTGAAGVGLGWLMDRGRKISLL
CTVPTGFAVITTAARDVMMLGSRLLTGLAGVASLVAVPYISEITVPAARGLGSC
VOLMVVTGILLAVAGVLEWRMLAVLGCVPPTLMLLMCTMPTPRPLTHQIOERA
MAALRLMGSEEGEERPPVCAEHOGLAMLRPRGVRLPLTIGCLWVFOOLSCVNAI
MFYANLTFEAKRKDSLSAVYGIIOVLETAVALIMDRGRKLLALSGVIVFSM
SAFGTYFKLQSPGSPNSHGLVPLISAEPADVHLGLAVLGVSMCLPIAGFAVGWGP
IPWLMSLEIPLHIKGVATGCVLTNMFMAFLVTKFENSIWEILRPYGAFLTAAFCI
LSVLFITLFEVPRKGRITLQITAHFEGR"
BASE COUNT 355 a 673 c 580 g 479 t
ORIGIN
```

```
alignment_scores:
Quality: 2057.00 Length: 455
Ratio: 4.633 Caps: 2
Percent Similarity: 97.582 Percent Identity: 85.714
```

```
alignment_block:
us-09-516-493-7 x RNO245935
Align seg 1/1 to: RNO245935 from: 1 to: 2087

1 ArgArgValPheLeuAlaIlePheAlaIleAlaLeuGlyProLeuSerPhe 17
|||||.....
99 CGCGGGGTCTTCTTGGCACTTTCGCGCGCGCCCTGGGCCGCTGACCTT 148
17 eglyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArg 34
|||||.....
149 CGGCTTTCGGCTTGGCTACAGCTCCCGCCATCCAGCTCGCGGCGCA 198
34 IAlaIleProAlaProArgLeuAspAspAlaIleAlaIleAlaSerTrpPheGly 50
:::|||||.....
199 CGCGACCCCGCTGGCCCTAGCCCTCGAGACACTCGCGGCTCTGTTGGG 248
51 AlaValIleThrLeuGlyAlaIleAlaIleGlyValLeuGlyTyrPhe 67
|||||.....
249 GCGCTCTGACCCCTGGCGCTGCGGCGAGGGGCGCTCTGGCGCGCTGCT 298
67 ValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerValProp 84
|||||.....
299 CTTGAGACCGTTCAGGGGCGCAAGCTGAGCTCTGCTTGCACCGTGCCT 348
84 heValAlaGlyPheAlaValIleThrAlaIleAlaIleAspValTrpMetLeu 100
|||||.....
349 TCGTGACCGGTTTTGGCTGATCATCAGCGCGCGCGGATGCTGATGCTG 398
101 LeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 117
|||||.....
399 CTCGAGAGCGCGCTCCCTCACCGGCTTACCTGCGGAGATCGCTCATAGT 448
117 IAlaIleProValTyrIleSerLeuIleAlaTyrProAlaValArgGlyLeuL 134
|||||.....
449 GGCACCGGTTATATTCGGAATTCGCTACCCAGCGCTTGAGAGACTGC 498
134 euGlySerCysValGlnLeuMetValValAlaGlyIleLeuLeuAlaTyr 150
|||||.....
499 TCGGCTCCTGTTGTCGAGCTGATGTTGTCACTGGCAATCCTTGGCTTAC 548
151 LeuAlaGlyTyrValLeuGlyTyrPArgTTrpLeuAlaValLeuGlyCysVa 167
|||||.....
549 GTGGCAGGCTGGGCTCCAGAAATGCGCTGCTGCTGCTGGCTGCTGTGT 598
167 IPropSerLeuMetLeuLeuLeuMetCysPheMetProGluThrProAla 184
|||||.....
599 GCCCCCGACCTCATGCTGCTGCTCATGCTCATATGCCGAGACCCAC 648
184 rPheLeuLeuThrGlnHisArgArgGlnIleAlaIle...AlaLeuArg 199
|||||.....
649 GCTTCTCCTGTCAGCAACACAGTACAGAGGCGCAATGGCTGCCCTGCG 698
200 PheLeuTrpGlySerGlnGlyGlnTyrTrpGluAspProIleGlyAlaGlu 216
|||||.....
699 TTCTCTGGGCGCTCTGAGAGGCTGGAGAGAGCCCTTGTGGGCTGGA 748
216 u...GlnSerPheHisLeuAlaLeuLeuArg...ProGlyIleLeuTyrSp 232
|||||.....
749 GCACACAGGGCTTCAGCTGCTATGCTGAGAGCGCGCTGGGTGTCCACAAG 798
232 robHelIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyVal 248
|||||.....
799 CCTCATCATCGGCAATTTGCTCATGCTTTCACAGCAGCTGTGAGGGGTC 848
249 AsnAlaValMetPheTyrAlaGluThrIlePheGluGlnAlaLysPheLys 265
|||||.....
849 AACGCATCATGTTCTATGCAACACCATCTTTCAGAGGCGCAAGTTTCAA 898
265 sasPserSerLeuAlaSerValValValGlyValIleGlyValLeuPheT 282
|||||.....
899 GGACAGCAGCGCTGCGCTGCTGCTACGTGGGCAATCATCCAGGTTCTGTCA 948
```



```

595 GCCCCCCACCTCATGCTGCTGCTCATGCTCATGCCCCGAAACCCAC 644
184 rPheLeuLeuThrGlnHisArgArgGlnGlnAlaIle...AlaLeuArg 139
645 GTTTTCTCTCTCACTCAACACCACTACAGAGAGCCATGGCTGCTTGGCG 654
200 PheLeuTrpGlySerGlnGlnGlyTrpGluAspProProlIleGlyAlaG 216
695 TTCCTGTGGGGCTGTGAGAGAGGCTGGAAGAGCCCTGTTGGGGCTGA 744
216 u...GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTrpLysP 232
745 GCACAGAGGCTTCCAGCTGCTGCTGAGGGCCCTGCTCATACAGGC 794
232 rPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyVal 248
795 CCTCATCTCCGCGATTTCCCTCATGCTTCCAGCAGCAGCTGTCAGGGCTC 844
249 AsnAlaValMetPheTrpAlaGlnThrIlePheGlnGlnAlaLysPheLy 265
845 AATGCTATCATGTTCTATGCAACAGCATCTTCGAGAGGCCAAGTTCAA 894
265 sAspSerSerLeuAlaSerValValGlyValIleGlnValLeuPheTr 282
895 GGACAGACAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 944
282 hrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeu 298
945 CTCCTGTGGGGCTGCTCATGCTCATGCTCATGCTCATGCTCATGCTCATG 994
299 ValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTr 315
995 GCCTTGTGGGGCTGCTCATGCTCATGCTCATGCTCATGCTCATGCTCATG 1044
315 rPheLeuLeuThrGlnGlnGlyProGlyLysSerSerHisValAlaIle 332
1045 CTTCAAACTGAGCCAGAGGCTCCCAAGACCTCTCCCAAGTGAAGGCTG 1093
332 eValAlaProValSerAlaGlnProValAlaSerValGlyLeuAlaTrp 348
1094 ..CTGCCCATCCGCGGAGAGCTGTGATGTCCCAAGTGGGAGCTGCTG 1141
349 LeuAlaValAlaGlySerMetCysLeuPheIleAlaGlyPheAlaValGly 365
1142 CTGGCTGTAGCAGCATGTGCTCTTCATGCTGCTGCTGCTGCTGCTGCTG 1191
365 pGlyProIleProTrpLeuLeuMetSerGlnIlePheProLeuHisVal 382
1192 GGGACCCATCCCTGCTGCTCATGTCAGAGATCTTCCTGCTGATGTC 1241
382 ySGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPhe 398
1242 AGCGGTGGCTACCGCATCTGTGCTCTCAACCACTGTTTCATGAGCTTT 1291
399 LeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTrpG 415
1292 CTAGTGACCAAGAGATTGACAGCGCATGAGATCTCAGACCCCTACG 1341
415 ValAheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrL 432
1342 TGGCTTCTGGCTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1391
432 euPheCysValProGlnThrLysGlyLysTrpLeuGlnGlnIlePheAla 448
1392 TCACCGTGTGCTGAGACTAAAGCAGACACTGTGACACAGTCACAGCC 1441
449 HisPheGluGlyArg 453
1442 CATTTGAGGAGAGA 1456
seq_name: qb_ro:AF232061
seq_documentation_block:

```

```

LOCUS AF232061 1843 bp mRNA ROD 23-JUN-2000
DEFINITION Mus musculus glucose transporter GLUT8 mRNA, complete cds.
ACCESSION AF232061
VERSION AF232061.1 GI:8671757
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1843)
Carayannopoulos,M.O., Chi,M.M., Cui,Y., Pingsterhaus,J.M.,
McKnight,R.A., Mueckler,M., Devaskar,S.U. and Moley,K.H.
GLUT8 is a glucose transporter responsible for insulin-stimulated
glucose uptake in the blastocyst
Proc. Natl. Acad. Sci. U.S.A. 97 (13), 7313-7318 (2000)
JOURNAL 20319023
MEDLINE 2 (bases 1 to 1843)
REFERENCE Moley,K.H., Carayannopoulos,M.O. and Cui,Y.
AUTHORS Direct Submission
TITLE Submitted (07-FEB-2000) OB/Gyn, Washington University, 4566 Scott
JOURNAL Ave, St. Louis, MO 63110, USA
FEATURES
source location/Qualifiers
1..1843
/organism="Mus musculus"
/strain="129"
/db_xref="taxon:10090"
1..48
49..1482
/codon_start=1
/product="glucose transporter GLUT8"
/protein_id="AAF78366.1"
/db_xref="GI:8671758"
/translation="MSPEDPOTPLRPPRPPRRVPLASFPAALGPIPSFPL
GYSSPAISPIRRTAPRALRGDMSRWCAVYVLAAGGIIIGMILDDSGRSLILL
CTVPYTGAVTAAADVMILGRLITLACGVASIVAPVISEATRPAYKGLDSC
VOLMVVGTLLVAVGWLEWRMLVVGCVPPPLMLLQVMPETRELTQHOYOA
MALRPLNGSEDEGMEPPVGAEGHGFOLALRRPGIKPLIIGISLWYQOLSGVNAI
MEYANSIFEAEKFDKDSLAVTGIIVFTAVAAALIMRAGRLLALSGVIMWSM
SAFGYFKLTOSLPNSNSHVGVLPIAAEPVDVQGLAIVLAVGSMCLFIAGFVGMCP
PWLIMSEIPLHVKGVATGICVLTWMPAFVLTTRKESSVEMELRPGAWLTAARCAL
SVLFTLYVPEPRKGRLLDQVTAHFEGR"
1483..1843
3'UTR BASE COUNT 291 a 609 c 529 g 410 t 4 others
ORIGIN
alignment_scores:
Quality: 2047.50 length: 455
Ratio: 4.622 Gaps: 3
Percent Similarity: 97.363 Percent Identity: 86.374
alignment_block:
US-09-516-493-7 x AF232061 ..
Align seg 1/1 to: AF232061 from: 1 to: 1843
1 ArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPh 17
118 CGCGGGCTCTTCCCTGCTTCTGCGCGCGCTGAGAGCCCTCAGCTT 167
17 eGlyPheAlaLeuGlyTrpSerSerProAlaIleProSerLeuGlnArg 34
168 CGGCTTGGCGCTCGGCTGAGAGTCCCGCCATCCCAAGCTGCGGCGCA 217
34 AlaIleProProAlaProArgLeuAspAspAlaAlaIleSerTrpPheGly 50
218 CGGACCCCGCGGCTGCTGCGAGACAATGCGGCTCTGCTGCTGGG 267
51 AlaValValThrLeuGlnGlyAlaAlaGlyValLeuGlnGlyTrpLe 67
268 GCGGCTGTGACCTGCGGCTGCTGCGAGGGGCACTACTGGGCGCTGCT 317
67 uValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerValProp 84

```



```

90 CGCGGGGCTTCCTCGCTTCCTTCGCGCGGCTGCGGACCCCTCAACTT 139
17 eglphealaleuGLYtYrSerSerProAlaIleProSerLeuGLArGa 34
|||||
140 CGGCTTCGGGGCTGCTACAGCTCCCGCCATCCCGCAGCTGCGGCGCA 189
34 laAlaProProAlaProArGLueAspAspAlaAlaIaSerTrpPheGLY 50
:::|||||
190 CGGACCCCGCGGCGCTGCGGACCAATGCGGCGCTCGGTTCGGG 239
51 AlaValAlaThrLeuGLYAlaAlaIaGLYValLeuGLYTrpLe 67
|||||
240 GCGGTGTACCTGGGCGCTGCGAGGGGCACTGCGGCGCTGCGT 289
67 uAlaSPaGLaGLYArGLYrLeuSerLeuLeuCySerValProp 84
1:::|||||
290 CCGGACCGCTTCAGGGCGCAAGCTGAGCTCTTGCTTCGACCGTGCCT 335
84 heValaIaGLYpheAlaValaIleThrAlaIaGLInaSPValTrpMeLeu 100
1:::|||||
340 TCGTCACTGGCTTGGCTGCATCAGCCGCGCGGATGTGTGATGCTG 389
101 LeuGLYGLYArGLeuleuThrGLYLeuAlaGLYAlaIaSerLeuVa 117
|||||
390 CTCGGAGGCGGCTCTCTCAACCGGCTTGGCTGCGGAGTCCCTCACTAGT 439
117 lAlaProValTYrIleSerGLuIleArTYrProAlaValArGLYLeuL 134
|||||
440 GGCACCGGCTTACATCTCGGAATGCCCTACCCAGCTGTCCGAGGACTGC 489
134 euGLYSerCySValGLInLeuMetValaValaGLYIleLeuLeuAlaTYr 150
|||||
490 TCGGCTCCTGTGTGAGCTGATGTGTGCTACCTGCACTCCTCGGCTAT 539
151 LeuAlaGLYTrpValaLeuGLUTrPArGTTrpLeuAlaValaLeuGLYCySVal 167
:::|||||
540 GTGGCAGGCTGCTCTAGAGTGGCTGCGCTGCGCTGCTGCTGCTGT 589
167 lProProSerLeuMetLeuLeuLeuMetCySPheMetProGLUTrProA 184
|||||
590 GCGCCCGCACCCTCATGCTGCTGCTCATGTGCTACATGCCGGAACCCAC 639
184 rGLPheLeuLeuThrGLInIlaArgArGLInIlaIle...AlaLeuArg 199
|||||
640 GTTTCCTCCTCACTCAACACACAGTACAGAGGSCATGCTGCTGCGC 689
200 PheLeuTrpGLYSerGLInGLInGLYTrpGLInaSPProProIleGLYAlaGL 216
|||||
690 TTCCTGTGGGCTCTGAGAGGGCTGGAAGAGCCCTGCTGGGCTCA 739
216 u...GInSerPheHISLeuAlaLeuLeuArgGLInProGLYIleTYrLYSP 232
1:::|||||
740 GCACGACGGGCTTCAGCTGCGCTGCTGAGGCGCCCTGGCATTTACAGC 789
232 rPheIleIleGLYValSerLeuMetAlaPheGLInGLInLeuSerGLYVal 248
|||||
790 CCTCATCATGCGCATTTCCCTCATGTCTTCACAGCAGCTGTCAGGGGTC 839
249 AsnAlaValaMetPheTYrAlaGLUTrIlePheGLInGLInaLYSPheLY 265
|||||
840 AATGCTATATGCTTATGTCACACAGCATTTTCAGAGGCCCAAGTTCAA 889
265 sAspSerSerLeuAlaSerValaValaGLYValaIleGLInValaLeuPheTr 282
|||||
890 GCACAGACAGCTGCGCTCGCTCACTGCGCATTAATCCAGGCTCTGTTCA 939
282 hrAlaValaIaAlaLeuIleMetAspArgAlaGLYArGLeuleuLeu 298
|||||
940 CTGCTGTGGCGGCTCATCATGACAGAGGAGGCGAAGGCTGCTG 989
299 ValLeuSerGLYValaValaMetValaPheSerThrSerAlaPheGLYAlaTY 315
:::|||||
990 GCGTGTGCGGCTGTGATCATGTGTTGACATGATGAGCCCTTGGTACCTA 1039

```

```

315 rPheLYSerLeuThrGLInGLYProGLYAsnSerSerHisValaIaIleS 332
|||||
1040 CTTCAAACTGACCCAGAGACCTCCCGCAGCACTCTCCACAGCTAGGCTG 1088
332 eAlaProValaSerAlaGLInProValaSPaSerValaGLYLeuAlaTrp 348
:::|||||
1089 .GTCGCATTCGCGGAGCGCTGTGATGTGTCAAAGTGGAGCTGCGCTGG 1136
349 LeuAlaValaGLYSerMetCySLeuPheIleAlaGLYIlePheAlaValaGLYTr 365
|||||
1137 CTGGCTGTAGGACAGATGTGCTCTTCATTCGTGCTTGGGCTGGGCTG 1186
365 pGLYProIleProTrpLeuLeuMetSerGLUlePheProLeuHisVala 382
|||||
1187 GGGACCATTCGCCCTGGCTCTCATGTCAGAGATCTTCCCTGATGTCA 1236
382 ySGLYValaIaThrGLYIleCySValaLeuThrAsnTrpLeuMetAlaPhe 398
1237 AGGCTGTGCTACCGCATCTGTCTCTCACCAACTGCTCATGCTGCTT 1286
399 LeuValaThrLYSGLYPheSerSerLeuMetGLUValaLeuArgProTYrGL 415
|||||
1287 CTAGTGACCAAGAGTTCAACAGCGTCAAGATGAGATGCTCAGACCTTACGG 1336
415 yAlaPheTrpLeuAlaSerAlaPheCySlePheSerValaLeuPheTrL 432
|||||
1337 TGCCTTGTGCTACCGCTGCTTCTGCGCTCAGTGTCTTATTCACAC 1386
432 eupheCySValaProGLUTrLYSGLYSThrLeuGLInIleThrAla 448
11
1387 TGACCGTTGTCCCTGAGACATAAAGCAGAGACTGTGGAACAGTCAAGCC 1436
449 HisPheGLYArg 453
1437 CATTTCGAGGAGACA 1451
seq_name: gb_fo:MMU245936
seq_documentation_block: 2072 bp mRNA ROD 18-FEB-2000
LOCUS MMU245936
DEFINITION Mus musculus mRNA for glucose transporter (GLUTX1 gene).
ACCESSION AJ245936
VERSION AJ245936.1 GI:7018606
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
JOURNAL
TITLE
AUTHORS
MEDLINE
REFERENCE
1 (bases 1 to 2072)
J. Biol. Chem. 275 (7), 4607-4612 (2000)
2 (bases 1 to 2072)
Ibberson, M.R.
Direct Submission
Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
CH-1005, SWITZERLAND
FEATURES
location/Qualifiers
1..2072
/organism="Mus musculus"
/db_xref="taxon:10090"
21..1454
/gene="GLUTX1"
21..1454
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CAB75719.1"
/db_xref="GI:7018607"

```

/translaction="MSPEDPOETPLRPPPEARPRGRNVELASPAALGFLNPFAL
 GYSSPAIPSLRTPALRILGDNAWSWGAATVTLGGMLIDRSRKLSTLL
 CTPEVPAFATITARDYVMMILGGRLTLGLAGVAASIVAPYISEIATVPAYRGLISG
 VOLMVTGILLAYVAVGVLEMBRLAVGCPPTMLLMCMPTPRLLHQHOXA
 MALRLFLMGSEBEPVGAHEHGFOLALLRRGTLTKPIITIGISLWFOOLSGVNAI
 MEVANSIPEEKAFKDDSSIASVTVGIIQVLEFVAVALIMDRAGRLLLASVIMVFSM
 SAFCTYFKLQSLPSSNSHVLVPIAABPDVQGLMALVAGSMCLFAGFAYVMGPI
 PMLMSEIFPLHVKGVATGICVLTNMFMAFLVTKKEFNSVMEMLRPYGAFWLTAFCAL
 SVLETLVVPETKGRLEQVTAHEEGR"

BASE COUNT 351 a 673 c 584 g 464 t
 ORIGIN

alignment_scores:
 Quality: 2041.50 Length: 455
 Ratio: 4.608 Gaps: 3
 Percent Similarity: 97.363 Percent Identity: 85.934

alignment_block:
 us-09-516-493-7 x MMU245936 ..

Align seg 1/1 to: MMU245936 from: 1 to: 2072

1 ArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPh 17
 |||||.....
 90 CGCGGGCTTCCTCGCTTCCTCGCGCGCTTCGGAGACCCCTCAACTT 139
 17 eGlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlyArg 34
 |||||.....
 140 CGGCTTCGGCGCTCGCTACAGCTCCCGCCATCCCACTCGCGCGCA 189
 34 AlaAlaProAlaProArgLeuAspAspAlaAlaAlaSerTrpPheGly 50
 ::|||.....
 190 CGGACCCCGCGCGCGCTCGAGACAAATCGCGCTTCGGTTCCGG 239
 51 AlaValAlaThrLeuGlyAlaAlaGlyValLeuGlyGlyTrpLe 67
 |||||.....
 240 GCGCTGTGTACCTGGCGCTCGTGGAGGGGCAATCTGGCGCGCTGCT 289
 67 uValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerValProp 84
 |::|.....
 290 CCGGACCTTCAGGGCGGCAAGCTGAGCTTCCTTCGCTCGACCGCTGCC 339
 84 heValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeu 100
 |||||.....
 340 TCGTCACTGGCTTGTGCTCATCACCGCGCGGAGTGTGTGATGCTG 389
 101 LeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 117
 |||||.....
 390 CTGGAGGGCGCTCTCTCACCGCGCTTCGCGGAGTGGCTCACTACT 439
 117 AlaProValTyrIleSerGlnIleAlaTyrProAlaValArgGlyLeuL 134
 |||||.....
 440 GGCACCGGTTTACATCTCGGAATCGCTACCCAGCTGTCCGAGGACTGC 489
 134 euGlySerCysValGlnLeuMetValValGlyIleLeuLeuAlaTyr 150
 |||||.....
 490 TCGGCTCCCTGTGTGAGCTGATGTGTGTCTGCGCATCTCTCGGCTAT 539
 151 LeuAlaGlyTrpValLeuGlnTrpArgTrpLeuAlaValLeuGlyCysVal 167
 ::|||.....
 540 GTGGAGGGCTGGGCTCTAGAGTGGCGCTGGCGCTGGGCTGTGT 589
 167 LProProSerLeuMetLeuLeuLeuMetCysPheMetProGluThrPro 184
 |||||.....
 590 GCCCCCGACCTTCATCTGCTGCTCATGTGTACATGCCCGAGACCCGCA 639
 184 rGlyPheLeuLeuThrGlnHisArgArgGlnGluAlaIle...AlaLeuArg 199
 |||||.....
 640 GTTTCTCTCTCACTCAACACGATCCAGAGGACCATGCTGCTGCGCC 689
 200 PheLeuTrpGlySerGluGlnGlyTrpGluAspPropTrpIleGlyAla 216
 |||||.....

690 TTCCTGTGGCGCTTCGAGGAGGCTGGGAAGACCCCTGTGGGCGTCA 739
 216 u...GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysP 232
 |.....
 740 GCACCAAGGCTTCACAGCTGCGCTGCGAGGCGCTTCGATCTACAAAC 789
 232 rPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyVal 248
 |||||.....
 790 CCTCTCATTCGGCATTTCCCTCATGCTCTTCACAGAGCTGTACAGGGCT 839
 249 AsnAlaValMetPheTyrAlaGluThrIlePheGluGlnAlaLysPhe 265
 |||||.....
 840 AATGCTATCATGTTCATGCAACAGCATCTTCAGAGGAGCCAGTTCAA 889
 265 SASPSerSerLeuAlaSerValValGlyValIleGlnValLeuPheT 282
 |||||.....
 890 GCACAGCAGCTTCGCTCGCTACAGTGGGCAATACACAGTCTCTTTTA 939
 282 hrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeu 298
 |||||.....
 940 CTGCTGTGGCGGCGCTCATCATGAGACAGAGGCGGAAGGCTGCTCTG 989
 299 ValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAla 315
 ::|||.....
 990 GCGCTTCGGGTGTGATCATGTGTTCATGATGAGTGCCTTGTGTAAC 1039
 315 rPheLysLeuThrGlnGlyGlyProGlyLysSerSerHisValAlaIle 332
 |||||.....
 1040 CTTCMAACCTACCCACAGAGCTCCCAAGCAACTCTCCCAAGTACGCTG 1088
 332 eAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrp 348
 ::|||.....
 1089 ..GTGCCATTCGCGGCGGAGCTGTGATGTCCAAAGTGGGACTGGCTGG 1136
 349 LeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTr 365
 |||||.....
 1137 CTGGCTGTAGGAGCAATGTCTTCATTCATGCTGGCTTGGCGTGGGCTG 1186
 365 pGlyProIleProTrpLeuLeuMetSerGlnIlePheProLeuHisVal 382
 |||||.....
 1187 GGGACCCATCCCTGCTGCTCATGTCAGAGATCTTCCTCGATGTCA 1236
 382 yGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPhe 398
 |||||.....
 1237 AGCGTGTGCTACCGCATCTGTCTTCACCACTGATGTTCACTGCGCTTT 1286
 399 LeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrG 415
 |||||.....
 1287 CTAGTGACCAAGAGTTCAACAGCGTCATGAGATGCTAGACCCCTACGG 1336
 415 yAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrL 432
 |||||.....
 1337 TGCCTTGTGGCTACCGGCTTCGCGCTCTCATGTCATTCACAC 1386
 432 euPheCysValProGluThrLysGlyLysThrLeuGlnGlnIleThrAla 448
 |||||.....
 1387 TGACCGCTGCTCCCTGAGACTAAAGCAGAGACTGTGAACAGTCAACAGCC 1436
 449 HisPheGluGlyArg 453
 |||||.....
 1437 CATTTGAGAGGACGA 1451

seq_name: gb_to:AB033418

seq_documentation_block:

LOCUS AB033418 2189 bp mRNA ROD 11-APR-2000
 DEFINITION Rattus norvegicus glut8 mRNA for glucose transporter 8, complete
 cds.

ACCESSION AB033418

VERSION AB033418.1 GI:7592743

KEYWORDS glucose transporter 8.

SOURCE Rattus norvegicus cDNA to mRNA, clone_lib: testis.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 2189)

AUTHORS Ishibashi, K.

TITLE Molecular cloning of a new putative glucose transporter

JOURNAL Published only in Database (2000) In press

REFERENCE 2 (bases 1 to 2189)

AUTHORS Ishibashi, K.

TITLE Direct Submission

JOURNAL Submitted (08-OCT-1999) to the DDBJ/EMBL/Genbank databases. Kenichi Ishibashi, Jichi Medical School, Pharmacology; Minami-Kawachi, Kawachi, Tochigi 329-0498, Japan (E-mail: kishiba@jichi.ac.jp, Tel: 81-285-58-7326, Fax: 81-285-44-5541)

location/Ovalifera

1. 2189

/organism="Rattus norvegicus"

/db_xref="taxon:10116"

/clone_lib="testis"

109. 1548

/gene="glut8"

109. 1548

/gene="glut8"

/note="Glut8 has 12 transmembrane segments"

/codon_start=1

/product="glucose transporter 8"

/protein_id="BA094383.1"

/db_xref="GI:7592744"

/translation="MSPEDPOETOPILRSPGARPGRRVFLATFAALGFLSPFALGYSSPAIPSLRRTAPALRLGDTAASWFGAVVILGAAAGCAVLGGMLLDGRKRLSLLCVLPVPTGAVITTAARDVWMLGGRLTGLACCAVLAIVTSETIAYPAVRGLLSCVQLMVYTGILLAYVAGWVLEWMLAVLGGCVPLMLLMCYMPETPRFLTOHOYOEAMALRFIEEAKFDSLSASVTGIIQVFTVAALIMBAGSKLLILASGVIMFVSMSAFGYFKLTGSGSNSSHVGLIPIASAPADVHGLAMLAVSGMCLFTAGFAVWGCPPIPMILMSEIFPIKIVKATGVCYLTNEMPAFLVTKKFNIMELIRYGAFLTAATCCILISVRLTFLVPTKRTLEQITTAHEGR"

BASE COUNT 392 a 702 c 609 g 486 t

ORIGIN

alignment_scores:

Quality: 2040.50 Length: 456

Ratio: 4.596 Gaps: 3

Percent Similarity: 97.368 Percent Identity: 85.307

alignment_block:

US-09-516-493-7 x AB033418 ..

Align seg 1/1 to: AB033418 from: 1 to: 2189

1 ArgatgvalPheleuAlaIaIaPheAlaIaIaLeuGIYProleuSerph 17

178 CGCCGGGCTTCTTGCCACCTTCGCCCGCCCTGGCGCCCTCAGCTT 227

17 eglYpheaIaLeuGIYTrSerSerProAlaIleProSerLeuGIhargA 34

228 CGGCTTGGCGGTGGCTACAGCTCCCGCCCATGCCAGCGCGGCGCA 277

34 IaaIaProProAlaProArgLeuAspAspAlaIaIaIaSerTrpPheGI 50

278 CCGCACCCCTGCCTACCGCTCGAGACACACGCGCCCTCGTTCCGG 327

51 AlayaIaIaThrLeuGIYAAlaIaIa...GIYGIYAlaIleuGIYTr 66

328 GCCGCTGACCCCTGGCGCTGGCGCCAGGCGCGCTGGCGCGCTG 377

66 pleuValAsparGalaIaIaArgIysLeuSerLeuLeuLeuCysSerValP 83

378 GCTCCTGACCGCTGACGCGCGCAAGCTGAGCTCTGCTGACCGCTG 427

83 ropheValaIaGIYpheaIaIaIleThrAlaIaIaGlnAspValTrpPheC 99

428 CCTTCGTGACCGGTTTTGCTGTATCATCCCGCGCGGATGTGTGATG 477

100 leuLeuGIYIaYArgLeuLeuThrGIYLeuAlaCysGIYValAlaSerle 116

478 CTGCTGGAGCGCGCCCTCCACCGCGGCTAGCGCTGGAGATCGGCTCAT 527

116 uValAlaProValTYrIleSerGIYIleAlaTYrProAlaValArgGIYL 133

528 AGTGCACCGGCTATATATCTCGGAATCGCCCTACCGCGCTTGCAGAG 577

133 euleuGIYSerCysValGlnIleMetValValValGIYIleLeuLeuAla 149

578 TGGCTGGCTCTGTGTGACCTGATGTTGTCACTGGCATCTCTTGGCC 627

150 TYrLeuAlaGIYTrPValLeuGIYTrPArgTrPLeuAlaValleuGIYCY 166

628 TAGTGGCAGCGTGGGCTCCAGATGGCGCTGGTGGCTGTGCTGGCGCTG 677

166 sValProProSerLeuMetLeuLeuLeuMetCysPheMetProGIYThrP 183

678 TGTGCCCCCGCACCCCTCATGCTGTGCTCATGTCTACATGCCGAGACCC 727

183 roArpPheLeuLeuThrGlnHisArgArgGlnIaIaIe...AlaLeu 198

728 CAGCTTCTCTCTGACGCAACACAGTACAGAGAGGCAATGAGCTGCTG 777

199 ArgPheLeuTrpGIYSerGIYGIYTrPGLuAspProIleGIYAl 215

778 CGCTTCTCTGTGGGCTCTGAGAGAGCGTGGAGAGCGCCCTGTGGGCG 827

215 aglu...GlnSerPheHisLeuAlaLeuLeuArgGlnProGIYIleTYrL 231

828 TGAGCACACAGCGCTTCAGCTGCTATGCTGAGCGCGCTGTGTCCACA 877

231 ysrProheIleIleGIYValSerLeuMetAlaPheGlnGlnIleuSerGIY 247

878 AGCCCTCATCATGCGCATTTGCTCATGTGCTTCCAGCAGCTGTCCAGG 927

248 ValAsnAlaValMetPheTYrAlaGIYThrIlePheGlnIuAlaIaIysPh 264

928 GTCAAGCGCCATCATGTCTTATGCCAACACCATCTTTAGAGAGGCAAGTT 977

264 elYAspSerSerLeuAlaSerValValValGIYValIleGlnValLeuP 281

978 CAAGGACAGCAGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1027

281 heThrAlaValAlaIaIaIleuIleMetAspArgAlaGIYArgLeuLeu 297

1028 TCACGTCTGTGGCGGCTCATGTGATGAGACAGAGCGAATAATGCTTT 1077

298 LeuValLeuSerGIYValValMetValPheSerThrSerAlaPheGIYAl 314

1078 CTGGCTTGTGGGCTGTGATCATGTGTTTTCAGCATGAGCCCTTTGGTAC 1127

314 aTYrPheIysLeuThrGlnGIYGIYProGIYAsnSerSerHisValAlaI 331

1128 CTACCTTCAAACTGACCCAGAGTGGGCCACCACTCTCCCATGTAGGCC 1177

331 IeSerIaProValSerAlaGlnProValAspAlaSerValGIYLeuAla 347

1178 TCCGTGGGCCATCTCCGACAGACCTGCTGATGTTCACCGGGGCTGGCC 1227

348 TrPLeuAlaValaGIYSerMetCysLeuPheIleAlaGIYpheaIaIaGI 364

1228 TGGCTGGCTGTAGGACAGATGTGCTCTTATGTGCTGTGTTTGCAGTAGG 1277

364 YTrpGIYProIleProTrPLeuLeuMetSerGIYIlePheProLeuHis 381

1278 CTGGGAGCCATCCCTGCTGCTCATGTGAGATCTTCCCTGTCAGACA 1327

381 alIysGIYValaIaIaThrGIYIleCysValaLeuThrAsnTrPLeuMetAla 397

1328 TCAAGGCTGTGGCTACCGGCTGTGTCTCTACCAACTGTTTCATGGCC 1377

```

398 PheLeuValThrIysGluPheSerLeuMetGluValIleuArgProTyr 414
|||||
1378 TTCTGTGTCACCAAGAGTTTACACAGATCATGAGATCTCTCAGACCTTA 1427
414 rGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValIleuPheT 431
|||||
1428 CGGCGCTTCCTGGCTCAGCCCTGCTTGTATCTCAGCGCTCTTTTCA 1477
431 hrLeuPheCysValProGluThrIysGlyIysThrLeuGluGlnIleThr 447
|||||
1478 CGCTCAGCTTTGTCTCCTGAGACTTAAGGACAGACTCTGGAACAATCACA 1527
448 AlaHisPheGluGlyArg 453
|||||
1528 GCCCATTTTCGAGGACGG 1545

seq_name: gb_cem:AF321324

seq_documentation_block:
LOCUS AF321324 1012 bp mRNA MAM 02-JUL-2001
DEFINITION Bos taurus glucose transporter 8 (Glut8) mRNA, partial cds.
ACCESSION AF321324
VERSION AF321324.1 GI:14582715
KEYWORDS
SOURCE
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 1012)
Augustin, R., Navarrete-Santos, A. and Fischer, B.
REFERENCE
Submitted (13-NOV-2000) Anatomy and Cell Biology, Faculty of
Medicine, Martin Luther University Halle, Grosse Steinstrasse 52,
Halle 061097, Germany
FEATURES
source
1..1012
/organism="Bos taurus"
/db_xref="taxon:9913"
<1..1012
/gene="Glut8"
<1..1005
/gene="Glut8"
/codon_start=1
/product="glucose transporter 8"
/protein_id="AAK69606.1"
/db_xref="GI:14582716"
/translation="ISELAYPEVRGILGSCVOLMVVGTILAYLAGNVLERNMLAVL
GVAPSEFMLIMCFEMDETPRELLSOKHDEAMAMQFMGICAGWHEEPVUGAHDHFV
AOLRRGVYKPRITIGISIMAFQOLSGVANVMRYATITIEAKFKDSSLASVVGIVY
LETTAAALIMDAGRRLILTLASGVVWFSTAFGYFKLTEGPNSSSHVDLPALVSM
EADADNVGILAMLVGNMCLIFAGFAVWGMPFPMILMSEIFPLHVGAVATGCVLTNMF
MAFLVTKFESSLMEVLRPYGAFWLASAFICFVGVLITLACVETKTKTLQJTAHREGR
"
BASE COUNT 153 a 323 c 302 g 234 t
ORIGIN
alignment_scores:
Quality: 1519.00 Length: 334
Ratio: 4.732 Gaps: 2
Percent Similarity: 96.108 Percent Identity: 87.126

alignment_block:
US-09-516-493-7 x AF321324 ..
Align seg 1/1 to: AF321324 from: 1 to: 1012

```

```

138 IGIuLeuMetValValIleGlyIleLeuLeuAlaTyrIleuAlaGlyTrpY 155
|||||
51 GAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 100
155 allLeuGluTrpArgTrpLeuAlaValLeuGlyCysValProSerLeu 171
|||||
101 TCGTCGAGTGGCGCTGGCGGCGTGGCGGCTGTGTGTGTGTGTGTGTGTGT 150
172 MetLeuLeuLeuMetCysPheMetProGluThrProAlaGluPheLeuTh 188
|||||
151 ATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 200
188 rGlnHisArgArgGlnIleAlaIle...AlaLeuArgPheLeuTrpGlys 204
|||||
201 TCAGCACAAGCACACGAGGACCATGCGCGCATGCGCATGCGCATGCGCTG 250
204 erGluGlnGlyTrpGluAspProIleGlyAlaGlu...GlnSerPhe 219
|||||
251 ATGCCAGAGGTTGGAGAGGCTCCCTTGGGGCTCAGCACAGGACCTTC 300
220 HisLeuAlaLeuLeuArgGlnProGlyIleTyrIysProPheIleIleG 236
|||||
301 CACGTGGCCAGCTGGGCGCTCTGTGTGTGTCAAGACGCTTCATCATCGG 350
236 yValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetP 253
|||||
351 CATCTCCTGATGGCTTCCAGCAGCTGTCAAGGCGTCAATGCCGTATGT 400
253 heTyrAlaGluThrIlePheGluGlnAlaLysPheLysAspSerSerLeu 269
|||||
401 TCTACGGGAGACCATCTTTGAGGAGGCTAAGTTCAAGACAGAGAGCTG 450
270 AlaSerValValIleGlyValIleGlnValLeuPheThrAlaValAla 286
|||||
451 GCTTCAGTGTGCTGGGCGCTATCAGAGCTGTCTGCTACTGCCAGCGGCG 500
286 AleuIleMetAspArgAlaGlyArgLeuLeuLeuValLeuSerGly 303
|||||
501 CTTGATCATGACAGAGCTGGCGAAGCTGTCTTGCATCTTGTCAAGTG 550
303 alValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLeuThr 319
|||||
551 TGTTCATGTGTGTTCAGACACAGCGCTTGTGTACACTTCAACGCGAC 600
320 GlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal 336
|||||
601 GAGGCGGCGCCAGCAACTCTCTCCAGCTGGACCTCCGCGCTGCTCTC 650
336 rAlaGlnProValAspAlaSerValIleLeuAlaTrpLeuAlaValIle 353
|||||
651 CATGAGAGCGCGCGATACCAAGCTGGGCTGGCTGGCTGGCTGGCTGGCT 700
353 erMetCysLeuPheIleAlaGlyPheAlaValIleGlyTrpGlyProIlePro 369
|||||
701 ACATGTGCTCTTCAACGCGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 750
370 TyrLeuLeuMetSerGluIlePheProLeuHisValLysGlyValAlaTh 386
|||||
751 TGGCTCTCATGTCAAGATCTTCCCTGTGCAATGTCAAAGGCGTGGCAC 800
386 rGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeuValTrpLysG 403
|||||
801 CGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 850
403 IuPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeu 419
|||||
851 AATTAGAGTCTCATGAGGTCGTCAGGCGCTAGCGCTGCTGCTGCTGCTG 900
420 AlaSerAlaPheCysIlePheSerValLeuPheTrpLeuPheCysValPr 436
|||||
901 GCCTGTGCTTCTGTGATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 950
436 ogIuThrLysGlyIysThrLeuGluGlnIleThrAlaHisPheGluGly 453

```

|||||
951 TGAGACCAAGGAGAGCTTGAGACAATACAGCCATTGTGAGGCC 1000
453 rg 453
11
1001 GA 1002

seq_name: gb_pat:AX191507

seq_documentation_block:

LOCUS AX191507 1461 bp DNA PAT 15-AUG-2001
DEFINITION Sequence 29 from Patent WO0149728.
ACCESSION AX191507
VERSION AX191507.1 GI:15209697
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1461)
AUTHORS Kato,S. and Kimura,T.
TITLE Human proteins having hydrophobic domains and dnas encoding these proteins
JOURNAL Patent: WO 0149728-A 29-12-JUL-2001;
Proteome Inc. (JP) : SAGAMI CHEMICAL RESEARCH CENTER (JP)
FEATURES
source Location/Qualifiers
1..1461
/organism="Homo sapiens"
/db_xref="taxon:9606"
61..849
/note="unamed protein product"
/codon_start=1
/protein_id="CAC51153.1"
/db_xref="GI:15209698"
/translation="MTPEPPEPTPLGPGGSPARGRVFLAFAAALGPSFGFAL
GYSPATPSIORAPPARLDDAASWFGAVYTUGAAGVYLGLVDRAGKLSLL
CSVPFAGFAVITAAQDWMLLGGRLTGLACGASLVAIPYISLTAIPAVAGLLGSC
VOLNVMVIGILLAYLAGWVLEWMLAVLGVCPSSLMLMCEMPEPRLTTOHROEA
APGLVRCGHGVLRLRLQADPQMPQLLARGHLAGLCTAC"

BASE COUNT 195 a 501 c 456 g 309 t
ORIGIN

alignment_scores:
Quality: 1427.00 Length: 455
Ratio: 4.725 Gaps: 4
Percent Similarity: 66.374 Percent Identity: 66.154

alignment_block:

US-09-516-493-7 x AX191507 ..
Align seg 1/1 to: AX191507 from: 1 to: 1461

1 ArgGValPheLeuAlaIlePheAlaIleAlaLeuGlyProLeuSerP 17
|||||
130 CGCGGGCTCTCTCGCGCGCTTGGCGCTTGGCGCTTGGCGCTT 179
17 eGlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlyArg 34
180 CGGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTT 229
34 AlaIleProAlaIleProAlaIleAspAspAlaIleAlaIleSerTrpPheGly 50
|||||
230 CGCGCGCGCGCGCGCGCGCTTGGCGCTTGGCGCTTGGCGCTT 279
51 AlaValValThrLeuGlyAlaIleAlaIleGlyValLeuGlyGlyTrpLe 67
|||||
280 GCTGCTGCTACCTTGGTGGCGCGCGCGGAGTCTTGGCGCGCTGCT 329
67 yValAspArgAlaGlyArgGlyLeuSerLeuLeuLeuCysSerValProp 84
|||||
330 GGTGGACCGCGCGCGCGCAAGCTTGTGCTGTGCTGCTGCTGCT 379

84 heValAlaGlyPheAlaValIleThrAlaIleAlaGlnAspValTrpMetLeu 100
|||||
380 TGTGGCGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTT 429
101 LeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 117
|||||
430 CTGGGGGGCGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCT 479
117 AlaProValTyrIleSerGlnIleAlaTyrProAlaValArgGlyLeuL 134
|||||
480 GGCGCGGCTTACATCTCGGAATCGCTACCGACAGCATCGCGGGGTGCG 529
134 euGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaTyr 150
|||||
530 TCGGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579
151 LeuAlaGlyTyrTrpValLeuGlnTyrArgTrpLeuAlaValLeuGlyCysVal 167
|||||
580 CTGGCAGGCTGGGTGCTGGAGTGGCGGCTGGCTGGCTGGCTGGCTGGCT 629
167 LProProSerLeuMetLeuLeuLeuMetCysPheMetProGlnTrpProA 184
|||||
630 GCCCGCTCTCCATGCTGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 679
184 rGpHeLeuLeuThrGlnHisArgArgGlnGlnAlaIleAlaLeuArgph 200
|||||
680 GCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 718
200 eLeuTrpGlySerGlnGlnGlyTrpGlnAspProIleGlyAlaGlnG 217
718 .. 718
217 InsThrPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrTrpProPhe 233
718 .. 718
234 IleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAl 250
718 .. 718
250 aValMetPheTyrAlaGlnThrIlePheGlnGlnAlaIleAlaIlePheLysAspS 267
718 .. 718
267 erSerLeuAlaSerValValValGlyValIleGlnValLeuPheThrAla 283
718 .. 718
284 ValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuValLe 300
719 .. 718
300 uSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheL 317
727 GTCAGGCTGTGTCATGTGTTCACACAGAGTGCCTTGGCGCTTACCTTCA 776
317 yLeuLeuThrGlnGlyGlyProGlyLysSerSerHisValAlaIleSerAla 333
777 AGCTGACCGAGGCTGGCGCTTGGCAACTCTGCGACGTGGCCATCTCGGCG 826
334 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTyrPLeuAl 350
827 CTTGCTCTGTGACAGCTGTGTGATGCCAGCTGGGGGTGCGCTGCTGCG 876
350 aValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTyrPLe 366
877 CTTGGGCGAGCATGTGCTTTCATGCGCGGAGGTC 911
367 ProIleProTrpLeuLeuMetSerGlnIlePheProLeuHisValLysG 383
911 .. 911
383 yValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeuV 400


```

KEYWORDS      .
SOURCE         Rattus sp.
ORGANISM       Rattus sp.
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                Rattus.
REFERENCE      1 (bases 1 to 2011)
AUTHORS        Thorens,B., Ibberson,M. and Oldry,M.
TITLE          Glutx polypeptide family and nucleic acids encoding same
JOURNAL        Patent: WO 0104145-A 13 18 -JAN-2001;
                University of Lausanne (CH)
FEATURES
  source        1..2011
                /organism="Rattus sp."
                /db_xref="taxon:10118"
                /note="unnamed protein product"
  CDS           /codon_start=1
                /protein_id="CAC28501.1"
                /db_xref="GI:12711211"
                /translation="MOEPLRAEGLDYDFPEAPASPEEKARAGALONRRVFLATFPA
                VLGNSFGALVYTSPIVLPALKRSSDPALRLDKIGMSVGVFTLGAAGGSAMLLN
                DLGKRKLSIMPSAVPSAIGYALMACARGIMLLIGMLTGNOGLTAACIPYVSPFLA
                PGGYAGALGATPOLMAVRSJSLYALGILLPMRLVAGEPVLVITLLSTPNSPR
                FLKSRDEEALQALILWRADSEVHWEFEQIDNVRQSSRVSWAEEMPRVYRPLLI
                TVLMRFLOOLGTPIILVYLQTFDSTSVLSQDAIVAGVRLSLVAIAAVTMDLA
                GRKVLVYSASIMFVANIILGLVQIVPRTLPNSTVAVITLGGTCQPPAAVNYLTL
                IPLATMLFIMGYAMGMPITWLMSEVPLARGVASGLCVLSMLTAFVLKRYFL
                AVNARGLQVPPFFSAICLSILFTQCCVPETRGRSLRDLAEFHTRMKSRP"
  BASE COUNT   337 a      621 c      593 g      459 t      1 others
  ORIGIN
alignment_scores
  Quality:     956.00      Length:     465
  Ratio:       2.828      Gaps:       7
  Percent Similarity: 72.688      Percent Identity: 45.591
alignment_block:
US-09-516-493-7 x AX076679
Align seg 1/1 to: AX076679 from: 1 to: 2011
1 ArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPh 17
|||||
109 AGAAGGCTGTCCTCGCCACCTTCGCGCTGCGGCAATTTCAGCTT 158
17 eGlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArg 34
|||||
159 TGGGTATGCGCTGCTACACATCTCCAGTCATCCCTGCGCTGAAGCCCT 208
34 IaAlaProAlaProArgLeuAspAspAlaAlaAlaIleSerTyrPheGly 50
|||||
209 CTTCTGACCCAGCACTACGCGCTGGACAATAATCCAGGCATCTCGGTTGGC 258
51 AlaValAlaThrLeuGlyAlaAlaAlaGlyValLeuGlyGlyTyrPLe 67
|||||
259 TCCGTGTTCACCTCGGCTGCGCGCGCTGGGGGCGCTCAGTCTATGCTACT 308
67 uValAspArgAlaGlyArgGlyLeuSerLeuLeuLeuLeuLeuSerValProp 84
|||||
309 CAATGACCTCTTGGGCGGGAAGCTCACACATCATGTTTCCGCTGCCCT 358
84 heValAlaGlyPheAlaValAlaIleThrAlaAlaGlnAspValTyrPheLeu 100
|||||
359 CGGCGCATGGCTATGCACCTCATGCTGCTGCGCTGCGCTTGGATGCTT 408
101 LeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaIleSerLeuVal 117
|||||
409 CTGCTGGGAGATGCTGACAGCGCTTGGCGGGGAGACCTGCTGCTGCTG 458
117 IaIleProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuL 134
|||||

```

```

459 CATCCCGGTGACGTCGTGAGATTGCAACCCCGGGGTCTTCCGGGGGCC 508
134 eUGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaTyr 150
|||||
509 TGGGGGCCACGCCGACGACGATGCGCGCTGTGGATCCTGTCGTCTCTAT 558
151 LeuAlaGlyTyrValLeuGluTyrPArgTyrPLeuAlaValaLeuGlyCysVal 167
|||||
559 GCCCTGGCTCTTCTGCTGCTGCTGGCGATGGCTTGGCGCGCGGAGAGG 608
167 IProProSerLeuMetLeuLeuLeuMetCysPheMetProGluTyrThrPro 184
|||||
609 GCCTTCTGCTGATGATTCCTGCTGCTTGAAGCCCAACCTGCGCT 658
184 rPheLeuLeuThrGlnIleHisArgArgGlnGluAlaAlaIleAlaLeuArgPhe 200
|||||
659 GCTTCTCTACTGTCCTAAGAACCCGGGATGAGGAAGCACCTGAG...GCGCTG 705
201 LeuTyrP.....GlySerGluGlnGlyTyrPgluAspProIleGly 214
|||||
706 ATCTGCTGTCGAGCGGACTCTGAGTCCACCTGGAGATTGGACAGATCCA 755
214 Y.....AlaGluGlnSerPheHisLeuAlaLeuLeuArg.....G 226
756 GGACAACGTGCGGAGACAGATGACCGAGTGTGTGGCGGAGCCCTGGG 805
226 InProGlyIleTyrLysProPheIleIleGlyValSerLeuMetAlaPhe 242
|||||
806 AGCCCGCTGTGACCGGCTATTTCATTCATTACAGTGTGTGATGCTTTTCTG 855
243 GlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIlePhe 259
|||||
856 CAGCAGCTGACAGGACATCATCCCATCCGTCGTACTACAGACCATCTT 905
259 eGluGluAlaLys...PheLysAspSerSerLeuAlaSerValValValG 275
906 GCACGACAGCTCCGTGTGTCGCTCCCTCCGACGACGATGAGCTATAGTGTG 955
275 IyValIleGlnValLeuPheThrAlaValAlaAlaLeuLeuMetAspArg 291
|||||
956 GTGCTGTGAGGCTCCTGCTGTGTGCTGATTCGCGCGCTGTCCACATGACATC 1005
292 AlaGlyArgArgLeuLeuValLeuSerGlyValValMetValPheSer 308
|||||
1006 GCGGCGCGCAAAAGCTCCTGCTATGTCGATGTCACGATCATATGTTGTGCG 1055
308 rThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGly.....GlyP 323
1056 CAACCTGACGCTGGGGGCTGTATGTCACGCTTGTGCGCAAGACTGTGACCC 1105
323 rGlyAsnSerSerHisValAlaIleSerAlaProValSerAlaGlnPro 339
|||||
1106 CCAACAGCAGCTGTGAGATTCGTGACGCTTGGGGGACGACGAGACGCCCA 1155
340 ValAspAlaSerValGlyLeuAlaTyrPLeuAlaValaGlySerMetCysLe 356
|||||
1156 GCGCGGCGCTTCAACTATCTCACCTGATACCCGTCGTCGACCATGCT 1205
356 uPheIleAlaGlyPheAlaValaGlyTyrPgluProIleProTyrPLeuLeuM 373
1206 CTTCATTTATGGGCTATGCCATGCGTGGGGGCCCTTGTGACCTGCTCTCTA 1255
373 eLseGluIlePheProLeuHisValLysGlyValAlaIleThrGlyIleCys 389
|||||
1256 TGTCTGAGGTTCTGCGCCGCGCTGCGCGCGTGTGGCTCAGAGGCTCTGCG 1305
390 ValLeuThrAsnTyrPLeuMetAlaPheLeuValIThrLysGluPheSerSe 406
|||||
1306 GTGCTGTGACGCTGAGCTCACAGCCTTTCGTCACATAAGTACTTCTCT...C 1354
406 rLeuMetGluValLeuArgProTyrGlyAlaPheThrPLeuAla...SerAla 422
|||||
1355 TGGCAGTGAATGCTTGGGCTTCACAGGTGCTTCTTCTTCTTCTGCGGCC 1404

```

```

423 pheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrIy 439
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1405 ATCTGCCTGCTCAGCCTGCTTCACACAGGCTGCTGTGCTGAGACAG 1454
439 sGIyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 453
:::|||||:::|||||:::|||||:::|||||:::|||||:::
1455 GGGCCGCTCGCTGAGCAGATCGAGGCTTCTTCACACTCGC 1497

```


CC disorder, dementia, depression, epilepsy, ischaemic cerebrovascular
 CC disease, stroke, cerebral neoplasms, Pick's disease, Huntington's
 CC disease and Parkinson's disease, demyelinating diseases, mental disorders
 CC including mood, anxiety, Schizophrenia and seasonal affective disorder,
 CC muscle disorder including cardiomyopathy, myocarditis, polymyositis,
 CC dermatomyositis, arylthimias and asthma and immunological disorders
 CC including AIDS, adult respiratory distress syndrome (ARDS), allergies,
 CC anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
 CC syndrome, systemic lupus erythematosus and other diseases including
 CC sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary
 CC artery stenosis, Grave's disease, Cushing's disease, Addison's disease,
 CC glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers
 CC portiasis and viral, bacterial, fungal, helminthic and protozoal
 CC infections. TRICH DNA is useful in gene therapy and in diagnostic
 CC purposes.

XX
 SO Sequence 2080 BP; 300 A; 688 C; 632 G; 460 T; 0 other;

alignment_scores:

Quality: 2310.50 Length: 454
 Ratio: 5.100 Gaps: 1
 Percent Similarity: 99.780 Percent Identity: 99.559

alignment_block:

05-09-516-493-7 x AAD09552 ..

Align seg 1/1 to: AAD09552 from: 1 to: 2080

1 ArgArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPh 17
 115 CGCGCCCTCTTCTCGCGCCTTGGCGCTGCGCCCTGCGCCACACACTT 164
 17 eGlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArg 34
 165 CGGCTTGGCGCTGCGCTACAGCTCCCGGCCATCCCTAGCTGAGCGCG 214
 34 laAlaProProAlaProAlaProAlaAspAlaAlaAlaSerTyrPheGly 50
 215 CCGCGCCCG 264
 51 AlAlaValThrLeuGlyAlaAlaAlaGlyValLeuGlyTyrPhe 67
 265 GCGTCTGACCTGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 314
 67 uValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerValProp 84
 315 GGTGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 364
 84 heValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTyrPheLeu 100
 365 TCGTGGCGCGCTTGGCGCTATCACCGCGCGCGCGCGCGCGCGCGCGCT 414
 101 LeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 117
 415 CTGGGGGGCGCGCTCTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCT 464
 117 laAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeu 134
 465 GCGCGCGCGCTACATCCGAAATCGCTACACGAGTCGCGGGGCTGCG 514
 134 euGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaTyr 150
 515 TCGGCTCTCTGTGCAAGCTAAATGTCGTGTCGCGCATCTCTCTGGCTAC 564
 151 LeuAlaGlyTyrPheValLeuGlyTyrPheValLeuAlaValLeuGlyCysVal 167
 565 CTGGCAGGCTGGGTGCTGAGTGGCGCTGCTGCTGCTGCTGCTGCTGCT 614
 167 lProProSerLeuMetLeuLeuLeuMetCysPheMetProGlnTyrProAla 184
 615 GCCCGCCCTCCCTCAGCTGCTCTCATGCTTCAATGCCCGAGACCGCGCG 664

184 rghPheLeuLeuThrGlnHisArgArgGlnGluAlaIle...AlaLeuArg 199
 665 GCTTCTCTGCTGACACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 714
 200 PheLeuTyrGlySerGlnGlnGlyTyrPheLysProProIleGlyAlaGly 216
 715 TTCTGTGGGGCTCCGACAGGCGTGGCGAAGACCCCGCATCGGGGCTGA 764
 216 uGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyTyrIleTyrProp 233
 765 GCAGAGCTTTACCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCT 814
 233 heIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsn 249
 815 TCATCATCGGCGCTCCCTGATGCGCTTCCAGACGCTGCGGGGCTCAC 864
 250 AlAlaValMetPheTyrAlaGluThrIlePheGlnGluAlaLysPheLys 266
 865 GCCGTATGTTCTATGCAAGACCATCTTGAAGAGCAAGTTCAAGGA 914
 266 pSerSerLeuAlaSerValValAlaGlyValIleGlnValLeuPheThr 283
 915 CAGCAGCTGGCGCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 964
 283 laValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuVal 299
 965 CTGTGGCGCTCTCATCATGACAGACAGAGCGGCGAGGCTGCTGCTGCT 1014
 300 LeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPh 316
 1015 TTGTGAGGTGTGTCATGATGTTGTTGACAGCAGATGCTTGGCGCTCTT 1064
 316 eLysLeuThrGlnGlyGlyProGlyLysSerSerHisValAlaIleSera 333
 1065 CAAGCTGACCCAGAGGTGGCGCTGGCAACTCTGCGACGTGGCATCTCG 1114
 333 laProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTyrPhe 349
 1115 CGCTGCTCTTGCACAGCTGTTGATGCCAGCGTGGCGCGCGCGCGCTG 1164
 350 AlAlaGlySerMetCysLeuPheIleAlaGlyPheAlaValAlaTyrPhe 366
 1165 GCGTGGGAGCATGATGCTTTCATCGCGCGCGCTTGGCGTGGCGTGGCG 1214
 366 yProIleProTyrPheLeuMetSerGluIlePheProLeuHisValLysG 383
 1215 GCCATGCCCTGGCTCTCATGTGACAGATCTTCCCTGATGTCAGAGG 1264
 383 lYValAlaThrGlyIleCysValLeuThrAsnTyrPheLeuMetAlaPheLeu 399
 1265 GCGTGGCGACAGGCATCTGCTCTCACCACACTGGCTCATGGCTTTCTC 1314
 400 ValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 416
 1315 GTGACCAAGAGTTCAGCAGCTCATGAGGTCTCAAGGCCCTATGAGGC 1364
 416 aPheTyrPheAlaSerAlaPheCysIlePheSerValLeuPheThrLeuP 433
 1365 CTTCGTGGCTTGGCTGCGCTTCTGATCTTCAGTGTCTCTTCACTTGT 1414
 433 heCysValProGlnTyrLysGlyLysThrLeuGlnGluIleThrAlaHis 449
 1415 TCTGTGTCCCTGAACCTAAAGGAAAGACTCTGGAACAATACACAGCCAT 1464
 450 PheGlnGlyArg 453
 1465 TTTGAGGGGCGA 1476

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.AAF55865

seq_documentation_block:

ID AAF55865 standard: cDNA; 2217 BP.

XX

```

AC  AAF55865;
XX
XX  17-APR-2001 (first entry)
XX
XX  Human GLUTX1 coding sequence.
DE
XX  Human: GLUTX; gene therapy; vaccine; hexose transport modulator;
KM  hexose transport disorder; ischaemia; diabetes; hypoglycaemia; ss;
KW  hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
XX  Homo sapiens.
OS
XX  WO200104145-A2.
PN
XX  18-JAN-2001.
PD
XX
XX  14-JUL-2000; 2000MO-1B01042.
PF
XX
XX  14-JUL-1999; 9905-0143907.
PR  27-AUG-1999; 9905-0151140.
PR  23-FEB-2000; 2000US-0184285.
PR  13-JUL-2000; 2000US-0616132.
XX
XX  (UYLA-) UNIV LAUSANNE.
XX
XX  Thorens B, Ibberson M, Uldry M;
PI
XX  WPI: 2001-112615/12.
DR  P-PSDB; AAB66932.
XX
XX  Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT  the prevention, diagnosis and treatment of hexose transport disorders,
PT  e.g. ischemia and diabetes.
XX
XX  Claim 3: Page 70-71; 124pp; English.
PS
XX
XX  The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC  AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
CC  glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC  function. The GLUTX proteins may be used in the diagnosis, prevention and
CC  treatment of hexose transport disorders such as ischaemia, diabetes,
CC  hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC  neurodegenerative disease. The present sequence is the coding sequence
CC  for human GLUTX1.
XX
XX  Sequence 2217 BP; 320 A; 750 C; 672 G; 475 T; 0 other:
SQ

```

alignment_scores:

Quality: 2301.50 Length: 454
 Ratio: 5.092 Gaps: 1
 Percent Similarity: 99.559 Percent Identity: 99.119

alignment_block:

US-09-516-493-7 x AAF55865

Align seg 1/1 to: AAF55865 from: 1 to: 2217

```

1  ArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPh 17
   |||||
417 CGCGCGCTTCTTCGCGCGCTTCCGCTCCCTGCGCGCCACCTCAGCTT 466
   |||||
17  eGlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnAla 34
   |||||
467 CGGCTTCCGCTTCGCTTACAGCTCCCGGCACTCCTAGCCTGACGCGG 516
   |||||
34  laAlaProAlaProArgLeuAspAlaAlaAlaSerTrpPheGly 50
   |||||
517 CCGGCGCCCGCGCCGCGCGCGGACGACGCGCGCGCTTCGCTTCCG 566
   |||||
51  AlaValAlaThrLeuGlyAlaAlaAlaGlyValLeuGlyGlyTrpLe 67
   |||||
567 GCTGTGCTGACCTTCGCTGCGCGCGCGGAGTGTGCGCGCTTGGCTT 616

```

```

67  uValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerValProp 84
   |||||
617 GGTGACCGCGCGCGCGGCAAGCTGAGCCTTGTGCTGCTCGGTGCGCT 666
   |||||
84  heValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeu 100
   |||||
667 TCGTGCCTCCGCTTTCGCGTCATCCACCGCGCCGACGAGCGTGTGATGCTG 716
   |||||
101  LeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 117
   |||||
717 CTGGGGGGCGCGCTTCTTCACCGGCTGGCGCTGGGTTGGCTCCCTAGT 766
   |||||
117  lAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeu 134
   |||||
767 GGCGCCGCTTCATCATCTCCGAATTCGCTAACCCAGCAGTCGGGGGCTTGC 816
   |||||
134  euGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaTyr 150
   |||||
817 TCGGCTCTGTGTGTCAGCTAATGTGCTGCTCGGCACTCCCTGGGCTTAC 866
   |||||
151  LeuAlaGlyTrpValLeuGluTrpPargTrpLeuAlaValLeuGlyCysVa 167
   |||||
867 CTGGCAGGCTGGGTGCTGGAGTGGCGCTGGCTGGCTGGGCTGGCTG 916
   |||||
167  lProProSerLeuMetLeuLeuLeuMetCysPheMetProGluThrPro 184
   |||||
917 GCCCGCTCCCTCATCTGCTGCTTTCATGTGCTTCATGCGCCGAGACCCGC 966
   |||||
184  rgPheLeuLeuThrGlnIleArgArgGlnGlnAlaIle...AlaLeuArg 199
   |||||
967 GCTTCCTGCTGACTGACGACAGGCGCCAGAGCCATGTGGCGCGCTTCGG 1016
   |||||
200  PheLeuTrpGlySerGluGlnGlyTrpGluAspProTrpIleGlyAlaG 216
   |||||
1017 TTCCTGTGGGGCTCCGAGAGGGCTGGAAAGACCCCGCATTCGGGGCTGT 1066
   |||||
216  uGlnSerPheIleSleAlaLeuLeuArgGlnProGlyIleTyrLysProp 233
   |||||
1067 GCAGAGCTTTCACCTGCGCTGCTGGCGGACGCCGCGCATTCACAGCCCT 1116
   |||||
233  heIleGlyValSerLeuMetAlaPheGlnIleLeuSerGlyValAsn 249
   |||||
1117 TCATCATCGCGCTTCCTGATGCGCTTCACAGCAGCTGTGGGGGTCAAC 1166
   |||||
250  AlaValMetPheTyrAlaGluThrIlePheGluGlnAlaLysPheLysAs 266
   |||||
1167 GCCGTATGTTCATGACAGACCATCTTGAAGAGCCCAAGTTCAAAGA 1216
   |||||
266  pSerSerLeuAlaSerValValValGlyValIleGlnValLeuPheThr 283
   |||||
1217 CAGCAGCTGCGCGCTGCGTGTGCTGGTGTCACTCAGGTGCTGTTCACAG 1266
   |||||
283  lAlaValAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuLeuVal 299
   |||||
1267 CTGTGGCGGCTCTCATCATGAGACAGAGCGGAGGCGTCCCTCGTGC 1316
   |||||
300  LeuSerGlyValValMetValPheSerThrSerAlaPheGlyValAlaTrp 316
   |||||
1317 TTTCTAGGCTGTGTGCTATGCTGTTCACACGACGAGCTTGGGCGCTTACT 1366
   |||||
316  euLysLeuThrGlnGlyIleProGlyAsnSerSerHisValAlaIleSer 333
   |||||
1367 CAAGCTGACCCAGGGTGGCGCTGGCAACTCTCTGCACAGTGGCCATCTCG 1416
   |||||
333  lAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeu 349
   |||||
1417 CGGCTGTCTGTGACAGCCTGTGATGCGCAGCGTGGGCGCTGGCTGCG 1466
   |||||
350  AlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValAlaGlyTrp 366
   |||||
1467 GCGTGTGGCAACATGTGCTTCATCGCGGCTTTCGCTGGTGGCGTGGG 1516

```

366 yProIleProTrpIleuMetSerGluIlePheProIleuHisValIysG 383
 1517 GCCCATCCCTGGCTCCTCATGTCTCAGACATCTCCCTCTGCATGTCAAGC 1566
 383 lValAlaIleThrGlyIleCysValIleuThrAsnTrpIleuMetAlaPheIeu 399
 1567 GCGTGGCAGCATTGCGTCTCTCACCACCTGCTCATGGCTTCCTTC 1616
 400 ValThrIysGluPheSerSerIleuMetGluValIleuArgProTyrGlyAl 416
 1617 GTGACCAAGAGATTGACAGCCTCATGAGGTCTCTCAGCCCTATGAGAC 1666
 416 alphaTrpIleuAlaSerAlaPheCysIlePheSerValIleuPheThrIeuP 433
 1667 CTTCGTGCTTGGCTCCTGCTTCCTCATCTTCAGTCTTTCACCTTGT 1716
 433 heCysValProGluThrIysGlyIleuThrIeuGluIleThrAlaHis 449
 1717 TCTGTCTCCCTGGAATTAAAGAAAGACTGTGAACAATTCACAGCCCAT 1766
 450 PheGluGlyArg 453
 1767 TTTCGAGGCGCGA 1778

seq_name: /stids2/gcgdata/geneseq/geneseqn/MA2001.DAT.AAF55866

seq_documentation_block:

ID AAF55866 standard; cDNA; 2087 BP.

AC AAF55866;

DT 17-APR-2001 (first entry)

DE Rat GLUTX1 coding sequence.

XX Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;

KM hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;

KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.

OS Rattus sp.

PN W0200104145-A2.

PD 18-JAN-2001.

PF 14-JUL-2000; 2000WO-IB01042.

PR 14-JUL-1999; 99US-0143907.

PR 27-AUG-1999; 99US-0151140.

PR 23-FEB-2000; 2000US-0184285.

PR 13-JUL-2000; 2000US-0616132.

PA (UTLA-) UNITV LAUSANNE.

PI Thorens B, Ibberson M, Uldry M;

DR WPI: 2001-112615/12.

DR P-PSDB: AAB66933.

XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in

PT the prevention, diagnosis and treatment of hexose transport disorders,

XX e.g. ischaemia and diabetes -

PS Claim 3; Page 71-73; 124pp: English.

XX The present invention relates to GLUTX proteins (AAF55865-AAF55871 and

CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative

CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport

CC function. The GLUTX proteins may be used in the diagnosis, prevention and

CC treatment of hexose transport disorders such as ischaemia, diabetes,

CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a

CC neurodegenerative disease. The present sequence is the coding sequence

CC for rat GLUTX1.

XX
 SQ Sequence 2087 BP; 355 A; 673 C; 580 G; 479 T; 0 other;

alignment_scores:
 Quality: 2057.00 Length: 455
 Ratio: 4.633 Gaps: 2
 Percent Similarity: 97.582 Percent Identity: 85.714

alignment_block:
 US-09-516-493-7 x AAF55866 ..

Align seq 1/1 to: AAF55866 from: 1 to: 2087

1 ArgArgValPheIleuAlaAlaPheAlaAlaAlaIleuGlyProIleuSerP 17
 99 CGCGGGGTCTTCTTGGCACTTCCGCGCCGCTGCGCCCTGCGGCTGCACTT 148
 17 eGlyPheAlaIleuGlyIleuSerSerProAlaIleProSerIleuGluArg 34
 149 CGGCTTGGCGCTTGGCTACAGCTCCCGCCCATCCCGAGCTGCGGCGCA 198
 34 lAlaIleProAlaIleProArgIleuAspAlaAlaAlaIleuSerTrpPheGly 50
 199 CGCGACCCCTGCTTACGCTTACGAGACACAGCTGCGGCTGCTGCTGCGG 248
 51 AlaValAlaThrIleuGlyAlaAlaIleuGlyValIleuGlyIleuPhe 67
 249 GCGGTCTGACCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 298
 67 uValAspArgAlaGlyArgIleuSerIleuIleuIleuIleuIleuIleuIleu 84
 299 CTGGAACCTGCGAGGCGGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTG 348
 84 heValAlaGlyPheAlaValAlaIleuAlaAlaIleuAlaIleuIleuIleu 100
 349 TCGTACCGGT 398
 101 IeugIyIyArgIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 117
 399 CTGGAAGCGCGCTCTCTACCGGCTTACGCTGCGGAGTGTGCTCATTAAGT 448
 117 lAlaIleuAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 134
 449 GGCACCGGT 498
 134 eugIySerCysValIleuIleuMetValValIleuIleuIleuIleuIleu 150
 499 TCGGCTCTGT 548
 151 IeulAlaGlyTrpValIleuIleuIleuIleuIleuIleuIleuIleuIleu 167
 549 GTGGCAGGCTGGGCTCTAGAAATGGCGCTGTGTGTGTGTGTGTGTGTGT 598
 167 lProProSerIleuMetIleuIleuMetCysPheMetProIleuIleuPro 184
 599 GCCCCACACCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 648
 184 tPheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 199
 649 GTTCTCTCTGACGACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 698
 200 PheIleuTrpGlySerIleuIleuIleuIleuIleuIleuIleuIleuIleu 216
 699 TTCTCTGTGGGCTGTGAGAGGCTGGGAAAGAGCCCTGTGTGGGCTGCA 748
 216 u...GlnSerPheHisIleuAlaIleuIleuIleuIleuIleuIleuIleu 232
 749 GCACGAGGCTTCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 798
 232 rPheIleuIleuGlyValSerIleuMetAlaIleuIleuIleuIleuIleu 248
 799 CCTCATCATCGCAATTGCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 848

```

249 AsnAlaValMetPheThrAlaGluThrIlePheGluGluAlaIlePheLeu 265
      |||.....|||.....|||.....|||.....|||.....|||.....
849 AACGCCATCATGTCATGCCAACACCATCTTGAGAGAGCCCAAGTTCAA 898
265 sAspSerLeuAlaSerValValAlaIleGluValLeuPheT 282
      |||.....|||.....|||.....|||.....|||.....|||.....
899 GCACAGACCCCTGGCTCCGTCACCTGGGACATCATGCCATCTCTTCA 948
282 hValAlaValAlaLeuIleMetAspArgAlaGlyArgArgLeuLeu 298
      |||.....|||.....|||.....|||.....|||.....|||.....
949 CTGCTGTGGGCCCCCATCATCATGACAGACAGGCGCAAAACGCTTCTG 998
299 ValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTy 315
      |||.....|||.....|||.....|||.....|||.....|||.....
999 GCGTTGTGGCGGTGATCATGCTTTTCACATGAGCGCCCTTGTCACCTA 1048
315 rPheIleuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleS 332
      |||.....|||.....|||.....|||.....|||.....|||.....
1049 CTTCAAACTGACCCAGAGTGGCCCGCAGCAACTCTCCATGTAGCCCTCC 1098
332 eAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrp 348
      |||.....|||.....|||.....|||.....|||.....|||.....
1099 TGGTCCCATCTCCGAGAGCCTGCTGATGTTCACTGGGCGCTGGCTGG 1148
349 LeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValAlaTy 365
      |||.....|||.....|||.....|||.....|||.....|||.....
1149 CTGGCTGTAGGAGCATGTGCTCTTCATCGCTGTTTCACATAGGCTG 1198
365 pGlyProIleProTrrPLeuLeuMetSerGluIlePheProLeuHisVal 382
      |||.....|||.....|||.....|||.....|||.....|||.....
1199 GGGACCCATCCCTGGCTCCATCATGACAGATCTTCCCTTCGACATCA 1248
382 ySGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPhe 398
      |||.....|||.....|||.....|||.....|||.....|||.....
1249 AAGGTGTGGTACCGGCTGTGTCTTACCAACTGTTTCATGAGGCTTT 1298
399 LeuValThrIleGlyPheSerSerLeuMetGluValLeuAlaArgProTy 415
      |||.....|||.....|||.....|||.....|||.....|||.....
1299 CTGGTACCAAGAGTTTACACATCATGAGATCTTCAGACCCCTACGG 1348
415 yAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrL 432
      |||.....|||.....|||.....|||.....|||.....|||.....
1349 CCGCTTCTGGCTACCGCGCTTCTGTATCCTCAACGCTCTTTCACGC 1398
432 euPheCysValProGluThrIleGlyIleThrLeuGluGlnIleThrAla 448
      |||.....|||.....|||.....|||.....|||.....|||.....
1399 TCACCTTGTCCCTGAGACTAAAGGAGGACTCTGGAACAATCACAGCC 1448
449 HisPheGluGlyArg 453
      |||.....|||.....|||.....|||.....|||.....|||.....
1449 CATTCGAGGCGAGG 1463
seq_name: /SIDS2/9c9data/geneseq/geneseq/NA2001.DAT:AAF55867
seq_documentation_block:
ID   AAF55867 standard; cDNA: 2072 bp.
XX
AC   AAF55867;
XX
DE   17-APR-2001 (first entry)
XX
DE   Murine GLUTX1 coding sequence.
XX
XX   Murine: GLUTX1; gene therapy; vaccine; hexose transport modulator;
KM   hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
KM   hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS   Mus sp.
XX
PN   W0200104145-A2.
XX
PD   18-JAN-2001.

```

```

XX 14-JUL-2000; 2000WO-IB01042.
PF 14-JUL-1999; 99US-0143907.
XX 27-AUG-1999; 99US-0151140.
PR 23-FEB-2000; 2000US-0184285.
XX 13-JUL-2000; 2000US-0616132.
XX (UyLA-) UNTV LAUSANNE.
XX PA
XX PI Thorens B, Ibberson M, Uldry M;
XX DR WPI: 2001-112615/12.
XX DR P-PSDB: AAB66934.
XX
PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
PT e.g. ischaemia and diabetes -
XX
XX PS Claim 3; Page 73-74; 124pp; English.
XX
CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischaemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is the coding sequence
CC for murine GLUTX1.
XX
SQ Sequence 2072 BP: 351 A; 673 C; 584 G; 464 T; 0 other:

alignment_scores:
Quality: 2041.50 Length: 455
Ratio: 4.608 Gaps: 3
Percent Similarity: 97.363 Percent Identity: 85.934

alignment_block:
US-09-516-493-7 x AAF55867 ..

Align seq 1/1 to: AAF55867 from: 1 to: 2072
1 ArgArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPh 17
|||.....|||.....|||.....|||.....|||.....|||.....
90 CGCGGGCTTCTCTCGCTTCTTCCGCGCCGCTGTGGAGACCCCTCACTT 139
17 eGlyPheAlaLeuGlyTyTrSerSerProAlaIleProSerLeuGlnArg 34
|||.....|||.....|||.....|||.....|||.....|||.....
140 CGGCTTCCGCTCGGCTACAGCTCCCGCCGATCCGACGCTCGGCGCA 189
34 lAlaIleProAlaIleProArgLeuAspAspAlaAlaAlaSerTrpPheGly 50
|||.....|||.....|||.....|||.....|||.....|||.....
190 CGGACCCCGCGCCCTCGCTCGAGACAATGCGGCTCTGCTTGGTGGG 239
51 AlAValAlaThrLeuGlyAlaAlaAlaGlyValLeuGlyGlyTrpLe 67
|||.....|||.....|||.....|||.....|||.....|||.....
240 GCGGCTGTGACCTGGGCGCTGCTGCGAGGGGCACTACTGGGCGCTGGCT 289
67 uValAspArgAlaGlyArgIleuSerLeuLeuLeuCysSerValProp 84
|||.....|||.....|||.....|||.....|||.....|||.....
290 CCTGAGACCTTTCAGAGGCGCAAGCTGAGCCCTTCTGTCACCGTCCCT 339
84 hValAlaGlyPheAlaValAlaIleThrAlaAlaGlnAspValTrpMetLeu 100
|||.....|||.....|||.....|||.....|||.....|||.....
340 TCGGATGTGGCTTGTGCTGATCACCGCGGCCGCGGATGTGTGATCTG 389
101 LeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 117
|||.....|||.....|||.....|||.....|||.....|||.....
390 CTGGAAGCCGCGCTCTACAGCGGCTTCTGCGGATCCGCTACTAGT 439
117 lAlaProValTyTrIleSerGluIleAlaTyTrProAlaValAlaGlyLeu 134
|||.....|||.....|||.....|||.....|||.....|||.....

```

440 GGCACCGTTTACATCTCGGAATACCGCTACCCAGCTGTGCGAGACATGCG 489
 134 euglysercysvalglnleumetvalvalvalglylleuleuvalatyr 150
 490 TCGGCTCTGCTGCGAGCTGATGTTGTACATCGGCATCTCTGCGCTAT 539
 151 leuvalaglytrpvalleuglutrpargrtrpleuvalalvalleuglycysva 167
 540 GTGGCAGGCTGGTCTAGAGCTGCGCTGCGCTGCTGGGCTGTGT 589
 167 lproproserleumetleuleumetcysphemetproglutnpro 184
 590 GCCCCCCCATCATGTGCTGCTCTCATGTGCTACATGCCGAGACCCGCC 639
 184 rghleuleuleutrhrghnlsargarglnglalaile...alaleuarg 199
 640 GTTTCTCTCTCTACTACACACGATACAGAGAGCCATGGCTCCCTGGCC 689
 200 pheleutrpglysergluglnglytrpgluaspproproilleglyalagl 216
 690 TTCTGTGGGCTGTAGAGAGGGCTGGGMAAGCCCTGTGGGGCTGA 739
 216 u...glnserphehlsleuleuarglnproglyletrylryp 232
 740 GCACCAAGGCTTCCAGCTGGCCCTCTGAGGCGCTGACATCTCAACAC 789
 232 ropheileleglyvalserleumetalapheglnleuenserglyal 248
 790 CCCCTCATCATCGGCATTTCCCTCATGTGCTTCCACAGCTGTGACGGGCT 839
 249 asnalavalmetphetrralagluthrillepheglnleualalysphely 265
 840 AATGTATCATGTCTATGCAACAGCATCTTCGAGAGAGCCCAAGTCA 889
 265 saspserserleualaservalvalvalglyalilleglnvalleuphet 282
 890 GGACAGCAGCCCTGGCTCGCTACGTGGGCAATTCAGGTCTGTCTCA 939
 282 hrvalavalalaleuilemetaspargalaglyargargyleuleu 298
 940 CTGCTGTGGCGGCTCTCATCATGAGACAGAGCGAGGCTGCTCTCTG 989
 299 valleuserglyvalvalmetvalpbeserthrseralaphelyalaty 315
 990 GCCTTGTGGGTGATCATGTGTTAGTATGATGCTTGTGGTACCTA 1039
 315 rpheylsleuthrlnglylproglyasnserserhlsvalalalles 332
 1040 CTTCAACTGACGAGCGCTTCCAGCAACCTCCACAGTACGAGGCTTG 1088
 332 eralaprovalseralaglnprovalaspalaservalglyleuvalatrp 348
 1089 ..CTGCCCATCGCGGAGAGCTGTGATGTCCAGTGGGACTGGGCTTG 1136
 349 leuvalavalglysermetcysleupheillealaglyphealavalglytr 365
 1137 CTGGCTGAGGAGCATGTGCTTTCATTTGCTTGGCTTGGGCTG 1186
 365 pglyproleuprottrpleuleumetsergluilepheproleuhsval 382
 1187 GGGACCCCATCCCTGGCTCTCATGTGAGATCTTCCCTGCAWGTCA 1236
 382 ysglvalalathrglylleucysvalleuthrAsntrpleumetlape 398
 1237 AGGGTGTGCTACCGGATCTGTCTCTCACCACATGGTTCATGGCTTT 1286
 399 leuvalthrlysgluPheserSerleumetgluvalleuargprotyr 415
 1287 CTAGAGACCAAGAGTTCACACAGCTCATGAGATGCTCAGACCTACGG 1336
 415 yalapherrpleualaseralaphecysllepheservalleuphetrl 432
 1337 TGCCTTGTGGCTCACCGCTTCTGCGCTCTCAGTGTCTCATTTACAC 1386

432 eupheCysValaProgluThrlysglyLysThrleuGluGlnIleThrAla 448
 11 |||||||||||||||||||||||||||||||||||||||||||||||
 1387 TCACCGTGTGCTCCCTGAGACTAAAGCAGAGACTCTCGAACAGTACACACC 1436
 449 HisPheGluGlyArg 453
 1437 CATTTGAGAGGAGCA 1451
 seq_name: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT:AA12574
 seq_documentation_block:
 ID AAD12574 standard; cDNA; 1461 BP.
 AC AAD12574;
 XX
 DT 25-SEP-2001 (first entry)
 XX
 DE Human protein having hydrophobic domain encoding cDNA clone HP10784.
 KW Human: hydrophobic domain; gene therapy; nutritional supplement;
 KW cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;
 KW multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;
 KW haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;
 KW Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;
 KW haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
 KW contraceptive; antiinfectility; antiinflammatory; ss.
 XX
 OS Homo sapiens.
 XX
 FH Location/Qualifiers
 FT CDS 61..849
 FT /*tag= a
 FT /product= "Human protein having hydrophobic domain"
 FT /note= "CDS is specifically is claimed in claim 3"
 FT
 PN WO200149728-A2.
 XX
 PD 12-JUL-2001.
 XX
 PE 28-DEC-2000; 2000MO-JP09359.
 XX
 PR 06-JAN-2000; 2000JP-0000585.
 PR 06-JAN-2000; 2000JP-0000588.
 PR 11-JAN-2000; 2000JP-0002299.
 PR 03-FEB-2000; 2000JP-0026862.
 PR 03-MAR-2000; 2000JP-0058367.
 XX
 PA (PROT-) PROTEGENE INC.
 PA (SAGA) SAGAMI CHEM RES. CENT.
 XX
 PI Kato S, Kimura T;
 DR WPI: 2001-418355/44.
 DR P-PSDB: AAE06579.
 XX
 PT Human proteins with hydrophobic domains and the nucleic acids encoding
 PT them, useful for preventing diagnosing and treating e.g. cancer,
 PT Alzheimer's and inflammation -
 PS
 XX
 PS Claim 4: Page 287-289; 563pp; English.
 XX
 CC The present sequence is human protein with hydrophobic domain encoding
 CC cDNA clone HP10784. The polynucleotide and polypeptide of the invention
 CC may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate polypeptide expression. The polynucleotides
 CC may be used to produce the polypeptide, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. The
 CC polynucleotides and its complementary sequences may also be used as DNA
 CC probes in diagnostic assays and also used in gene therapy. The
 CC polypeptides may also be used as antigens in the production of antibodies
 CC and in assays to identify modulators of polypeptide expression and
 CC activity. The polypeptides and nucleic acids may be used as nutritional

CC supplements, to modulate cytokine and cell proliferation activity, to
 CC modulate immune stimulation or suppression (e.g. for the treatment of
 CC microbial infections and autoimmune disorders such as multiple sclerosis,
 CC rheumatoid arthritis and insulin-dependent diabetes); to modulate
 CC hematopoiesis, to modulate tissue growth activity (e.g. for the
 CC treatment of Parkinson's disease, Huntington's disease and Alzheimer's
 CC disease), to modulate activin and inhibin activity (e.g. for controlling
 CC fertility), to modulate chemotactic and chemokinetic activity, to
 CC modulate haemostatic and thrombolytic activity, to modulate receptor
 CC ligand activity, to modulate inflammation and to inhibit tumour growth.
 XX
 SQ Sequence 1461 BP; 195 A; 501 C; 456 G; 309 T; 0 other:

alignment_scores:

Quality: 1427.00 Length: 455
 Ratio: 4.725 Gaps: 4
 Percent Similarity: 66.374 Percent Identity: 66.154

alignment_block:

us-09-516-493-7 x AAD12574 ..

Align seg 1/1 to: AAD12574 from: 1 to: 1461

1 ArgArGValPheLeuAlaIaPheAlaIaIaLeuGlyProLeuSerPh 17
 |||||
 130 CGCGCGCTTCCTCGCGCGCTTCGCGCGCTCGCGCGCACTGACGCTT 179
 17 eGlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArg 34
 |||||
 180 CGGCTTCGCGCTGCGCTACACCTCCCGCGCATCCCTACGCTGCACCGCG 229
 34 IaAlaProProAlaProArGLeuAspAspAlaIaIaIaSerTrpPheGly 50
 |||||
 230 CGCGCGCGCGCGCGCGCGCGCTGACAGACGCGCGCGCTCGTTCGGG 279
 51 AlaValIaIaThrLeuGlyAlaIaIaIaGlyIaIaLeuGlyIaIaTrpPhe 67
 |||||
 280 GCTTCGTCGACCTGGTGCGCGCGCGCGGAGTGCTGGCGCGCTGGCT 329
 67 uValAspArgAlaGlyArgGlyLeuSerLeuLeuLeuLeuGlySerValProp 84
 |||||
 330 GGTGGAACGCGCGCGCGCAAGCTGAGCTTCGCTGCTCGCGCGCT 379
 84 heValaIaGlyPheAlaValIaIaThrAlaIaIaGlnAspValaIaTrpMetLeu 100
 |||||
 380 TCGTGGCGCGCTTCGCGCTGACATCACCGCGCGCGCGCTGTCGATGCTG 429
 101 LeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValaIaIaSerLeuVa 117
 |||||
 430 CTGGGGGCGCGCTCTCACCGCGCGCTGGCTGGCTGGCTGCTGCTGCT 479
 117 IaIaProValaIaIaIaSerGluIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 134
 |||||
 480 GCGCGCGCTGACATCTCGAAATCCCTACCAAGCAGACGCGGGGCTTCC 529
 134 euGlySerCysValaIaIaLeuMetValaIaIaIaIaIaIaIaIaIaIaIaIa 150
 |||||
 530 TCGGCTCTGCTGTCAGCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 579
 151 LeuAlaGlyIa 167
 |||||
 580 CTGGCAGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 629
 167 IaProProSerLeuMetLeuLeuLeuLeuMetCysPheMetProGluThrPro 184
 |||||
 630 GCGCGCGCTCCCATGCTGCTTCATGCTTCATGCTTCATGCTTCATGCT 679
 184 rGpHeLeuLeuThrGlnHisArgArgGlnGlnIaIaIaIaIaIaIaIaIaIa 200
 |||||
 680 GCTTCCTGCTGACTGACGACGACGCGCGCGAGGAGGCTGCTC..... 718
 200 eleuTrpGlySerGlnGlyIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 217

718 718
 217 InSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIlePheProphe 233
 718 718
 234 IleIleGlyValaSerLeuMetAlaPheGlnGlnLeuSerGlyValaIaIa 250
 718 718
 250 aValaMetPheTyrAlaGluThrIlePheGluGluAlaIaLysPheLysAsp 267
 718 718
 267 erSerLeuAlaSerValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 283
 718 718
 284 ValaIaIaLeuIleMetAspArgAlaGlyArgArgLeuLeuVala 300
 719CTGGCTT 726
 300 uSerGlyValaIaIaMetValaPheSerThrSerAlaPheGlyAlaIaIaTrpPhe 317
 |||||
 727 GTCAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 776
 317 yLeuThrGlnGlnGlyIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 333
 |||||
 777 AGCTGACCCAGGCTGGCGCTGGCAACCTTCGACACGCGGCACTTCGCGG 826
 334 ProValaSerAlaGlnProValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 350
 |||||
 827 CCGTCTCTCTCACACGCTGTTGATGCCAGCGCTGGCGCTGGCGCTGGCG 876
 350 aValaIa 366
 |||||
 877 CCGGCGCAGCATGCTGCTTCATCGCGCGGAGGCTC..... 911
 367 ProIleProTrpLeuLeuMetSerGluIlePheProLeuHisValaIaIa 383
 911 911
 383 yValaIaIaThrGlyIleCysValaIaIaIaIaIaIaIaIaIaIaIaIaIa 400
 911 911
 400 aIa 416
 |||||
 912CTCAGCGCGCTATGAGGCC 929
 417 PheTrpLeuAlaIa 433
 |||||
 930 TTTCTGCTTCCTCCCTGCTTCATCTTCATCTTCATCTTCATCTTCATCT 979
 433 eCysValaIa 450
 |||||
 980 CTGTGCTCCGAAACTTAAGAAAGAACTCTGAAACAATGACACGCCATT 1029
 450 heGluGlyArg 453
 |||||
 1030 TTGAGGCGCGCA 1040
 seq_name: /SIDS7/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF55871
 seq_documentation_block:
 ID AAF55871 standard; cDNA; 2011 BP.
 XX AAF55871:
 AC
 AC
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE Rat GLUTX3 coding sequence.

```

XX      Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;
KM      hexose transport disorder; ischaemia; diabetes; hypoglycaemia; ss;
KM      hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS      Rattus sp.
XX
PN      WO200104145-A2.
XX
PD      18-JAN-2001.
XX
PF      14-JUL-2000; 2000WO-IB01042.
XX
PR      14-JUL-1999; 990S-0143907.
PR      27-AUG-1999; 99US-0151140.
PR      23-FEB-2000; 2000US-0184285.
PR      13-JUL-2000; 2000US-0616132.
XX
PA      (UYLA-) UNIV LAUSANNE.
XX
PI      Thorens B, Ibberson M, Uldry M;
XX
DR      MPI; 2001-112615/12.
XX
P-PSDB; AAB66938.
XX
PT      Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT      the prevention, diagnosis and treatment of hexose transport disorders,
XX      e.g. ischemia and diabetes -
XX
PS      Claim 3; Page 82-83; 124pp; English.
XX
CC      The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC      AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
CC      glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC      function. The GLUTX proteins may be used in the diagnosis, prevention and
CC      treatment of hexose transport disorders such as ischaemia, diabetes,
CC      hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC      neurodegenerative disease. The present sequence is the coding sequence
CC      for rat GLUTX3.
XX
SO      Sequence 2011 BP: 337 A; 621 C; 593 G; 459 T; 1 other;

alignment_scores:
      Quality: 956.00      Length: 465
      Ratio: 2.828      Gaps: 7
      Percent Similarity: 72.688      Percent Identity: 45.591

alignment_block:
US-09-516-493-7 x AAF55871 ..

Align seg 1/1 to: AAF55871 from: 1 to: 2011

1  AArgAGValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerph 17*
   |||||
109 AGAAGGTGTCTCTGCGCACCTTCGCTGCGCTGCTGGCAATTTCAGCTT 158
   |||||
17  eGlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArg 34
   |||||
159 TGGGTATGCTGCTGCTGCTACATCTCCAGTCATCCCTGCGCTGAAGGCT 208
   |||||
34  lAlaIleProAlaProArgLeuAspAspAlaAlaIleAspTTPheGly 50
   |||||
209 CTTTCACCCAGCAGCTACGCTGACAAAATCCAGGCAATCCGTTTGGG 258
   |||||
51  AlAlaValThrLeuGlyAlaAlaIleGlyValLeuGlyGlyTyrPhe 67
   |||||
259 TCCGTTTACCCCTGGGTGCGCGCTGGGGGCTCAGTGTCTACT 308
   |||||
67  uValAspAlaGlyArgLysLeuSerLeuLeuLeuCysSerValProp 84
   |||||
309 CATGACCTCTTGGGCGGAGAGCTCAGCATCATGTTTCCGCTGTCGCC 358

```

```

84  heValAlaGlyPheAlaValIleThrAlaAlaGlnAspValITrMetLeu 100
   ....
359 GGGCATTGGCTATGACCTACGCTGAGTGGCGGTGGCTGAGATGCTT 408
   |||||
101  LeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaIleSerLeuVa 117
   |||||
409 CTGCTGGGGAGAGATGCTACAGAGCTTTGGCCGGGAGACTCAGTCTGCTG 458
   |||||
117  lAlaProValTyrIleSerGlyIleAlaTyrProAlaValArgGlyLeuL 134
   |||||
459 CATCCCGGTGTACGCTGCTGAGATTGCACCCCGGTGTGCGGGGCGCC 508
   |||||
134  euGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaTyr 150
   |||||
509 TGGGGGCCAGCCCGAGCTCATGGCGGTGGATCCCTGCTCTCTAT 558
   |||||
151  LeuAlaGlyTyrValLeuGlnTyrPArgTyrPheAlaValLeuGlyCysVa 167
   |||||
559 GCCCTGGCTTCTGCTGCTGCGGATGCTTGGCGGGGAGAGG 608
   |||||
167  lProProSerLeuMetLeuLeuLeuMetCysPheMetProGlnThrProA 184
   |||||
609 GCTTCTGCTGCTATGATCCCTGCTGCTGCTATGCTCCTCAACCTCGCTC 658
   |||||
184  rPheLeuLeuThrGlnHisArgArgGlnGlnAlaIleAlaLeuArgPhe 200
   |||||
659 GCTTCTTACTGCTTACGAGCCGGAGTAGAGAGACTGAC...GCGCTG 705
   |||||
201  LeuITrP.....GlySerGlnGlnGlyTyrPLeuLysPropProIleG1 214
   |||||
706 ATCTGCTGCGAGCGGAGCTGAGCTGACCTGAGGATTTGAGCAGATCCA 755
   |||||
214  Y.....AlaGlnIleSerPheHisIleAlaLeuLeuArg.....G 226
   |||||
756 GGACACGTCGGGAGACAGATCCAGAGTCTGCGGAGCGGCTGAGG 805
   |||||
226  lProGlyIleTyrLysProPheIleIleGlyValSerLeuMetAlaPhe 242
   |||||
806 AGCCCGGTGTACCGGCTATTCTCATTTACAGTCTGATGCGTTTCTG 855
   |||||
243  GlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlnThrIlePh 259
   |||||
856 CAGACGCTGACAGGACATCCATCTCGTGTACCTACAGACATCTT 905
   |||||
259  eGlnGlnAlaLys...PheLysAspSerSerLeuAlaSerValValG 275
   |||||
906 CGACAGCAGTCCGTGCTGCCCTCCAGAGATGACGATTAATGTTG 955
   |||||
275  lValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetLaspArg 291
   |||||
956 GTGCTGTGAGGCTCCTGTCTGTCTGTATTCCTGCTGACACCATGAGCTG 1005
   |||||
292  AlAGlyArgArgLeuLeuValLeuSerGlyValValMetValPheSe 308
   |||||
1006 GCGCGCGGAAAGTCCGCTGTATGTGTGACATTCATATGTTTGGCG 1055
   |||||
308  rThrSerAlaPheGlyAlaLeuTyrPheLysLeuThrGlnGly.....GlyP 323
   |||||
1056 CAACCTGACGCTGGGCTGTATGTCCAGCTTGTGCCAAGGACTCTGACCC 1105
   |||||
1106 CCACAGCAGCTGTGAGATCGTACGCTTGGGGGCGAGCGAGCGCCCA 1155
   |||||
340  ValAspAlaSerValGlyLeuAlaTyrPLeuAlaValGlySerMetCys 356
   |||||
1156 GCGCGCGCTTCAACTATCTCACCCTGATACCCCTGCTG6CCACCATGCT 1205
   |||||
356  uPheIleAlaGlyPheAlaValAlaGlyTyrPLeuProIleProTyrPLeu 373
   |||||
1206 CTTTCATTATGGGTATGACCTGAGGCTGGGGGCGGACATCAGCTGCTCTTA 1255
   |||||
373  etSerGlnIlePheProLeuHisValLysGlyValAlaIleThrGlyLecys 389

```



```

1256 TGCTGAGTTCTTGCCCGCTGGCCCGCTGGTGGTGGCTTCACAGCGCTTGC 1305
390 ValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSer 406
1306 GTGCTGCTGACGGTGCCTCACAGCGCTTGCTGCTCACAAAGTACTTCTCTGC 1354
406 rLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeuAlaSerAla 422
1355 TGGCACTGAAATCCCTTGGCGCTCCAGAGTCCCTTTCTTCTTCTCGCGCC 1404
423 PheGlyIlePheSerValLeuPheThrIlePheGlyValProGluThrLys 439
1405 ATCTGCTGCTGACCGCTGCTTTCACAGCGTCTGTGTGCTGGACACAG 1455
439 sGlyLysThrLeuGluIleIleThrAlaHisPheGluGlyArg 453
1455 GGGCGCTGCTGGGAGCAATGAGGCGCTTTCACACACTGCG 1497

```

Quality:	928.00	Length:	474
Ratio:	2.737	Gaps:	11
Percent Similarity:	71.519	Percent Identity:	45.570
alignment block:			
US-09-516-493-7	x AAF55870	..	
Align seg 1/1	to: AAF55870	from: 1	to: 1541

PR	25-FEB-2000	2000EP-0301439	
PR	25-FEB-1999	9905-0121825	
PR	05-MAR-1999	9905-0123180	
PR	03-MAR-1999	9905-0123348	
PR	23-MAR-1999	9905-0125788	
PR	25-MAR-1999	9905-0126564	
PR	25-MAR-1999	9905-0126785	
PR	01-APR-1999	9905-0127462	
PR	06-APR-1999	9905-0128234	
PR	08-APR-1999	9905-0128714	
PR	16-APR-1999	9905-0129845	
PR	19-APR-1999	9905-0130077	
PR	21-APR-1999	9905-0130449	
PR	23-APR-1999	9905-0130510	
PR	23-APR-1999	9905-0130841	
PR	28-APR-1999	9905-0131449	
PR	30-APR-1999	9905-0132048	
PR	30-APR-1999	9905-0132407	
PR	04-MAY-1999	9905-0132484	
PR	05-MAY-1999	9905-0132485	
PR	06-MAY-1999	9905-0132486	
PR	06-MAY-1999	9905-0132487	
PR	07-MAY-1999	9905-0132683	
PR	11-MAY-1999	9905-0134256	
PR	14-MAY-1999	9905-0134218	
PR	14-MAY-1999	9905-0134221	
PR	14-MAY-1999	9905-0134219	
PR	14-MAY-1999	9905-0134370	
PR	18-MAY-1999	9905-0134746	
PR	18-MAY-1999	9905-0134741	
PR	20-MAY-1999	9905-0135124	
PR	21-MAY-1999	9905-0135353	
PR	24-MAY-1999	9905-0135629	
PR	25-MAY-1999	9905-0136021	
PR	27-MAY-1999	9905-0136392	
PR	28-MAY-1999	9905-0136782	
PR	01-JUN-1999	9905-0137222	
PR	03-JUN-1999	9905-0137528	
PR	04-JUN-1999	9905-0137502	
PR	07-JUN-1999	9905-0137724	
PR	08-JUN-1999	9905-0138094	
PR	10-JUN-1999	9905-0138540	
PR	10-JUN-1999	9905-0138647	
PR	14-JUN-1999	9905-0139119	
PR	16-JUN-1999	9905-0139452	
PR	16-JUN-1999	9905-0139453	
PR	17-JUN-1999	9905-0139492	
PR	18-JUN-1999	9905-0139454	
PR	18-JUN-1999	9905-0139455	
PR	18-JUN-1999	9905-0139456	
PR	18-JUN-1999	9905-0139457	
PR	18-JUN-1999	9905-0139458	
PR	18-JUN-1999	9905-0139459	
PR	18-JUN-1999	9905-0139460	
PR	18-JUN-1999	9905-0139461	
PR	18-JUN-1999	9905-0139462	
PR	18-JUN-1999	9905-0139463	
PR	18-JUN-1999	9905-0139464	
PR	28-JUN-1999	9905-0140623	
PR	28-JUN-1999	9905-0140625	
PR	28-JUN-1999	9905-0140921	
PR	30-JUN-1999	9905-0141267	
PR	01-JUL-1999	9905-0141842	
PR	02-JUL-1999	9905-0142151	
PR	02-JUL-1999	9905-0142055	
PR	06-JUL-1999	9905-0142393	
PR	08-JUL-1999	9905-0142803	

PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 26-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147935.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.

PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161922.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

alignment_scores: Quality: 657.50 Length: 469
Ratio: 2.192 Gaps: 11
Percent Similarity: 63.966 Percent Identity: 35.821

alignment_block:
US-09-516-493-7 x AAC36319 ..

Align seg 1/1 to: AAC36319 from: 1 to: 1741

3 ValPheLeuAlaIaIaPheAlaIaLeuGlyProLeuSerPheGlyPh 19
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
284 GTTATCTCAGCACATTCGTCGCGTCTGTCCTTTGCGTTCGCTTC 333
19 eAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlyArgAlaIaP 36
::: ||||||||||||||| || ||
334 TTGTGCGGATCTCTCACCCTCT.....CAGGCTGCATTTA 371
36 roProAlaProArgLeuAspAspAlaAlaIaIaSerTrpPheGlyAlaVal 52
||| ||| ||||||:||||:||||:||||:||||:||||:||||:||||
372 CGAATGACCTTTCATTGACATACCTGACGTTTCACCTTTCGTTCTT 421
53 ValThrLeuGlyAlaIaAlaIaGlyValLeuGlyGlyTrpLeuValAs 69
:::||||:||||:||||:||||:||||:||||:||||:||||:||||
422 CTAACTTTTGGCGCAATGATCGGTGCTATAACAACGCGCCTATACCTGA 471
69 parGAlaGlyArgLysLeuSerLeuLeuLeuGlySerValProPheValA 86
I ::|||:||||| ||||:||||:||||:||||:||||:||||:||||
472 TTTAGTTGGAAGAAGGGGCGATGAGAGTTTCCTCGCATTTTGTGTAG 521
86 IacIyPheAlaValIleThrAlaIaGlnAspValTrpMetLeuLeuGly 102
:::||||:||||:||||:||||:||||:||||:||||:||||:||||
522 TCGGCGGCTAGCAATCATCTTTGCCAAGGAGATGAGCTGTCGACCTC 571
103 GLyArgLeuLeuThrGlyLeuAlaIaCysGlyValAlaSerLeuValAlaPr 119
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

CC aberrantly located in the cytoplasm in breast cancer cells. In other
 CC cancers, cells can secrete excessive amounts of hormones e.g, cancers of
 CC the adrenal medulla can secrete excessive amounts of adrenaline and
 CC noradrenaline, leading to hypertension. TRAMP is expressed in cancer
 CC cells, and transport disorders result from either excessive or
 CC insufficient molecular transport. Anti-TRAMP antibodies and nucleic acids
 CC encoding TRAMP can be used as diagnostic tools for such disorders. TRAMP
 CC antagonists can be used to treat or prevent a cancer associated with
 CC increased TRAMP expression. Anti-TRAMP antibodies can be used directly
 CC as an antagonist or as a targeting mechanism for drugs. Alternatively,
 CC a TRAMP antisense nucleotide can be used to treat cancers. A TRAMP
 CC agonist or expression vector may be used to treat a disorder caused by
 CC reduced transport of biologically active molecules.

XX Sequence 1527 BP; 216 A; 519 C; 494 G; 298 T; 0 other;

alignment_scores:

Quality: 655.00 Length: 469
 Ratio: 2.266 Gaps: 17
 Percent Similarity: 61.620 Percent Identity: 39.232

alignment_block:

US-09-516-493-7 x AA211731 ..

Align seg 1/1 to: AA211731 from: 1 to: 1527

```

1 ArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPh 17
  ::::::::::::::::::::::::::::::::::::::::::::::::::::
129 AAAGAGGCTGTTCTGCGCCACCTTCCGCCAGTGTGCGCAATTTACGCTT 178
17 eGlyPheAlaLeuGlyTyrSerSerProAlaLeuProSerLeuGlnArg 34
  ::::::::::::::::::::
179 TGGGATGTCCTGTGCTACACATCCCTGTCAATCCAGCCCTGAGCGCT 228
34 LaAlaProAlaProArgLeuAspAspAlaAlaAlaSerTrpPheGly 50
  ::::::::::::::::::::
229 CTTGTGATCTGACCTGCATCTGACCAATCCAGGCAATCTGCTTGGG 278
51 AlaValAlaThrLeuGlyAlaAlaAlaGlyValLeuGlyAlaTyrPhe 67
  ::::::::::::::::::::::::::::::::::::::::::::::::::::
279 TCCGCTGTACACCTGCGAGCAGCGCGGAGGCTG.AGTGCCATGATCTT 327
67 uValAspArgAlaGlyArgLysLeuSerLeuLeuLeuLysSerValProp 84
  ::::::::::::::::::::::::::::::::::::::::::::::::::::
328 CAAGACCTCTCTGCGCGGAGAGCTGACATCATCTTTCACGCTGGCGCT 377
84 heValAlaGlyPheAlaValAlaLeuTrpAlaAlaGlnAspValTrpMetLeu 100
  ::::::::::::::::::::::::::::::::::::::::::::::::::::
378 CGGCGCGCGCTATGCGCTCATGCGGCTGCCACGCGCTTGAGATGCTG 427
101 LeuGlyAlaArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 117
  ::::::::::::::::::::::::::::::::::::::::::::::::::::
428 CTGCTGGAGAGAGCGCTGACGGGCTTCCGCGGGGCTTCACAGCTGCGCT 477
117 LaAlaProValTyrIleSerGlnIleAlaTyrProAlaValArgGlyLeuL 134
  ::::::::::::::::::::::::::::::::::::::::::::::::::::
478 CATCCGCTGCTAGCTGTGAGATGCTCCCGCAGCGCTTCTGGGGCTG 527
134 euGlySerCysValGlnLeuMetValValaGlyIleLeuLeuAlaTyr 150
  ::::::::::::::::::::::::::::::::::::::::::::::::::::
528 TGGGGGCGCACCCAGCTCATGGGAGTTCGCGATCCCTGCGCTTAC 577
151 LeuAlaGlyTrpValLeuGluTrrParGTrpLeuAlaValLeuGlyCysVal 167
  ::::::::::::::::::::::::::::::::::::::::::::::::::::
578 GCCCTTGGCGCTGCTGCTGCGGCGCTGCGCTGCTGCGGGGAGGCG 627
167 LProProSerLeuMetLeuLeuLeuMetCysPheMetProGluTrpProA 184
  ::::::::::::::::::::::::::::::::::::::::::::::::::::
628 GCCCTGCTCATCATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 677
184 rgPheLeuLeuThrGlnIleAspArgGlnGluAlaIle...AlaLeuArg 195
  ::::::::::::::::::::::::::::::::::::::::::::::::::::

```

```

678 GCTTCTGCTCTCTCGGGCAGGAGCAAGAGCCCTGCGGGGCTGCGCC 727
200 PheLeuTrpGlySerGlnGlnGly.....TrpGluAspProIleG1 214
  ::::::::::::::::::::
728 TGGCTGCGTGGAGCGGAGCTGCATGTCCACTGTGGAGTTTGAGACATGCCA 777
214 Y.....AlaGlnGlnSerPheHisLeu.....AlaLeuLeuArg 226
  ::::::::::::::::::::
778 GGACAAAGCTCCGAGAGACAGAGCAGCGAGTGTGGCTGAGAGCAGGG 827
226 LProGlyIleTyrLysProPheIleIleGlyValSerLeuMetAlaPhe 242
  ::::::::::::::::::::
828 CCCCACAGTGTGCGGCCATCACCGCTTGGCTGATGCGGCTTCTG 877
243 GlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluTrpLeuP 259
  ::::::::::::::::::::
878 CAGCAGCTGACGGGCGATCACGCCCATCTGCTGCTTACTGAGTCCATCTT 927
259 eGluGluAlaLysPhe.....LysAspSerSerLeuAlaSerV 272
  ::::::::::::::::::::
928 CGACAGACCGGCTGTCTGCTGCTGCGCCCAAGAGAGAC.....GCAG 968
272 AlValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIle 288
  ::::::::::::::::::::
969 CCATCTGTGGGCGCGTGGCGGCTCTGCTGCTGCTGATGCGCGCTTACC 1018
289 MetAspArgAlaGlyArgArgLeuLeuValLeuLeuSerGlyValAlaMe 305
  ::::::::::::::::::::
1019 ATGAGACTGCGCAGCGCGCAAGAGTGTCTGTTGCTGACAGCCACTGTC 1068
305 eValPheSer...ThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGly 321
  ::::::::::::::::::::
1069 TGTCCACAGCGGCGCATGATTTGTGCTGCC.....AACGTGACTTGGGG 1112
322 .....GlyProGlyAsnSerSerHisValAlaIleSerAl 333
  ::::::::::::::::::::
1113 CTGTATCATCACTTTTGGCCCGAGGCTTGAAGCCCAACAGCACTCGGG 1162
333 aproValSer.....AlaGlnProVal..... 340
  ::::::::::::::::::::
1163 CCGGAAAGCAGAGTCTGCGGGGAGCTTGGGCGAGCCCTTGACACCGCG 1212
341 ..... 344
1213 TGGCATCTACACCTGTGCTGCTGCTGCTGCGCACCATCTTTCATGATGG 1262
344 ValGlyLeuAla...TrpLeu.....AlaV 351
  ::::::::::::::::::::
1263 GTAAGTGTGTGTGTGTGTGCTGAGAGGGCAGAGCTGTGTGGGTTAAGGG 1312
351 aGlySerMet..... 354
  ::::::::::::::::::::
1313 TGGGTATGTGTGTGTGGGTGTGGTGTGAGAGGGGGTGTCCAGCAGCTCAGC 1362
355 .....CysIle 356
  ::::::::::::::::::::
1363 GGAGACAGACAGCGGCTTCCAGTCAACCCACAGGAGGCTGAACSTGCT 1412
356 urPheIle...AlaGlyPheAlaValGlyTrrGlyProIleProTrpLeuL 372
  ::::::::::::::::::::
1413 CCGCTCCAGCAGAGCTGACGCGTGGGCTGGGCTGATC.ACTTGAGTGC 1461
372 euMetSerGlnIlePheProLeuHisVal...LysGlyValAlaThrGlyL 388
  ::::::::::::::::::::
1462 TCATGCTGACAGTCTGCGCTGCG...GTGCGGTGGGTGCTGACAGGCT 1508
388 e 388
1509 A 1509
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AMC40459
seq_documentation_block:

```

ID AAC0459 standard; DNA: 1755 BP.
XX AAC0459;
AC
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 28375.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
protein identification; signal transduction pathway;
metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 08-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

```
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0156559.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160880.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.
```

```
alignment_scores:
  Quality: 653.00      Length: 460
  Ratio: 2.162        Gaps: 9
  Percent Similarity: 65.652  Percent Identity: 34.565
```

alignment_block:

US-09-516-493-7 x AAC40459 ..

Align seg 1/1 to: AAC40459 from: 1 to: 1755

```
3 ValPheLeuAlaIaPheAlaIaLeuGlyProLeuSerPheGlyPh 19
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
203 GTTCTCTCAGTACTTTGTTGCTGCTGCTGCTCTTTGAGCTTGGCTC 252
19 eAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaIaP 36
:|||||:|||||:|||||:|||||:|||||:|||||:
253 TTGTGTGGATACTCAGCGCCACTCACTCATATTAAGACAAGATCTTC 301
```

```
36 roProAlaIProArgLeuAspAspAlaIaIaSerTrpPheGlyAlaVal 52
:|||||:|||||:|||||:|||||:|||||:|||||:
302 .....AATCTTCCTCCAGAGTTCTCATGTTGGATCATC 340
53 ValThrLeuGlyValAlaIaIaGlyValLeuGlyGlyTrpLeuValAs 69
:|||||:|||||:|||||:|||||:|||||:|||||:
341 TTAACATACGTCGTCGAATGCTGCTGCTGTATGAGCGCAAAATTCAGA 390
69 ParGalIaGlyArgLysLeuSerLeuLeuLeuCysSerValProPheVal 86
:|||||:|||||:|||||:|||||:|||||:|||||:
391 TTTCTCCGCGCGCAAAAGGCGCATGAGAGCTCAGCTGCTTCGCATTA 440
86 IacGlyPheAlaValIleThrAlaIaGlnAspValTrpMetLeuLeuGly 102
:|||||:|||||:|||||:|||||:|||||:|||||:
441 CAGGTGGCTGCGCTGCTCTTCTTCCACCAAGGCGCATGTTACTTGAATGA 490
103 GlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaIleSerLeuValA 119
|||||:|||||:|||||:|||||:|||||:|||||:
491 GGAAGGTTCTTTACAGCAATATGCAATGGAGTTTCTTATGTGTGCC 540
119 oValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGly 136
|||||:|||||:|||||:|||||:|||||:|||||:
541 TGTGTACATTGCTGACATATCTCCCAAGATCTCCGAGGTGACCTACAA 590
136 ercysValGlnLeuMetValValIcGlyIleLeuLeuAlaTyrLeuAla 152
:|||||:|||||:|||||:|||||:|||||:|||||:
591 CACTGAACCAACTCATGATGTTGATCGGCTCATGCTTCTTCTTGATGC 640
153 GlyTrpValLeuGluTrpArgTrpLeuAlaValLeuGlyCysValProP 169
|||||:|||||:|||||:|||||:|||||:|||||:
641 GGATCTCTCATTTCTTGAAACACTGCTCCCTAACGCTACTCTCCCTG 690
169 oSerLeuMetLeuLeuLeuMetCysPheMetProGluThrProArgPhe 186
:|||||:|||||:|||||:|||||:|||||:|||||:
691 CATTTGTTTGTCTTTGGCTTGTGCTTTCATACCCGAATCTCCTCGATGG 740
186 euleu...ThrGlnHisArgArgGlnGlnAlaIleAlaLeuArgPheLeu 201
||:|||||:|||||:|||||:|||||:|||||:
741 TGCCAAAGAGCAGCCATGAGAAAGAGTTTCCGACCCCTGCAAAAGCTG 790
202 TrpGlySerGluGlnGlyTrpGluAspProprioGlyAlaGluGln 218
|||||:|||||:|||||:|||||:|||||:|||||:
791 CGAGCAAAAGATGCAGATATCACAAATGAAGCAGACGCTATTCAGACTC 840
218 rPheHis...LeuAlaLeuLeuArgGlnProGlyIleTyr..... 230
841 GATTCAAGCTCTAGACATTTCTCCAAAGCAACAATCCAAAGACTGTGT 890
231 .....LysProPheIleIleGlyValSerLeuMetAlaPhe 242
:|||||:|||||:|||||:|||||:|||||:|||||:
891 CCAAGAAATATATGTCGATCTGTCATCATTTGCTTCCCGATGATATTC 940
243 GlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIlePh 259
|||||:|||||:|||||:|||||:|||||:|||||:
941 CAACAGCTTTGTTGGAAATCAATGGATCGCAATTCACCAAGTAAACGTT 990
259 eGluGlnAlaLysPheLysAspSerSerLeuAlaSerValValValGly 275
|||||:|||||:|||||:|||||:|||||:|||||:
991 TGTAAAGCCGCAATTTACTCTGTGGAACACTGAGACAATGCTATCGCTT 1040
276 ValIleGlnValLeuPheThrAlaValAlaIaLeuIleMetAspArgAl 292
|||||:|||||:|||||:|||||:|||||:|||||:
1041 GTT...CAGTGCCCATTAACCTGTTCTGGAAACAATCTGTATACATTAATC 1087
292 acGlyArgArgLeuLeuValLeuSerGlyValValMetValPheSerTr 309
:|||||:|||||:|||||:|||||:|||||:|||||:
1088 TGGACGAAAGCCCACTAATATGATTTACGCTGCTGATATCTTCTGGGAT 1137
309 hisSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyGlyProGlyAsn 325
|||||:|||||:|||||:|||||:|||||:|||||:
1138 GCATTACTACAGGCAACATCTTTCTTACTCAAG..... 1169
326 serSerHisValAlaIleSerAlaIProValSerAlaGlnProValAspAl 342
```

1170GACA 1174
342 aSerValGlyLeuAlaTrp.....LeuAlaValGlySerMetCysL 356
1175 GAGCTTGCTGGTGAATGGGCTCCCTTCCCTTACCGCTGGAGGTGACTTA 1224
356 eupPheIleAlaGlyPheAlaValGlyTrpGlyProIleProIleProIleu 372
1225 TCTATGATGAGCTCTTCTCCATCGGAATGGACGCTGCTCCTGGGTGANA 1274
373 MetSerGluIlePheProLeuHisValLysGlyValAlaIleThrGlyTleCy 389
1275 ATGCTGAGATATTCCGATAAACGTAAGGAATCGCAGCAAGCTTAAGT 1324
389 sValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerS 406
1325 GGTACTAGGAATGTGCTGCTGGCTTCTTCTTACCTTCACT 1374
406 eTleuMetGluValLeuArpProTyrGlyAlaPheTrpLeuAlaSerAla 422
1375 TCCTCATGAGCTGGAGCTCTCCA...GGTACATTCTATTGTCAGCGCT 1421
423 PheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrLy 439
1422 TTTGCAGCTGCAGCATATATTGTGGCGAAGATGTCGACAGACGANA 1471
439 sGlyLysThrLeuGluGlnIleThrAla 448
1472 AGGGAAGACACTGGAAGACATCCAAGCT 1499
seq_name: /SIPS2/gcgdata/geneseq/geneseqn/NA2000.DAT: AAC39099
seq_documentation_block:
ID AAC39099 standard; DNA: 1683 BP.
AC AAC39099;
XX
XX AAC39099;
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23364.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KM Protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 16-APR-1999; 9905-0128714.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 28-APR-1999; 9905-0130891.
PR 30-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 04-MAY-1999; 9905-0132407.
PR 05-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.

PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139753.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144684.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.

PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151330.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:

Quality: 639.50 Length: 461
Ratio: 2.153 Gaps: 11
Percent Similarity: 64.425 Percent Identity: 35.792

alignment_block:

US-09-516-493-7 x AAC45857 ..

Align seg 1/1 to: AAC45857 from: 1 to: 1395

3 ValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPh 19
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
82 GPTTATCTCAGCACATTCGCTGCGGCTCTGCTTCTTTCGCTTCGCTTC 131
19 eAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaIaIap 36
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
132 TTCTGCGGCACTCTTTCACCTGCT.....CAGGCTGCATTA 169
36 roProAlaProArgLeuAspAspAlaAlaIaIaSerTrpPheGlyAlaIaVal 52

```

170 GGAATGACCTTCATGACTATAGCTGAGTTTCACTCTTGTTGTTCTTA 219
53 ValThrLeuGlyAlaAlaIleGlyValLeuGlyGlyTrpLeuValAs 69
220 CTAACCTTTGGCGCAATGATCGGTGCTATACACAGCGGCTATAGCTA 269
69 PArgAlaGlyArgLysLeuSerLeuLeuLeuLeuLeuLeuLeuLeu 86
270 TTTAGCTTGGAAGAAGGGCGATGAGATTCCCTCATCTTTGTGTAG 319
86 LagLPhaAlaValIleThrAlaIleGlnAspValTrpMetLeuLeu 102
320 TCGGCTGGCTACCAATCATCTTTGCCAAGGAGTAGTGCTCTGCACCT 369
103 GlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuAlaPr 119
370 GGAACATCGGCACAGCGGATATGGAAATGGAGCATTTTCTTATGTGTC 419
119 OValTyrIleSerGluIleAlaTyrProAlaValArgLysLeuLeu 136
420 AATCTTTATACAGAAATTTGCACCTAAACTTTTCAGAGGCTCTAACCA 469
136 eCys...ValGlnMetValValGlyIleLeuLeuAlaTyrLeu 151
470 CACGAAACACACAGATTTGATGCTGCACTGGAGTGTGCTTCTCATC 519
152 AlaGlyTyrValLeuGlyTyrArgTyrLeuAlaValLeuGlyCysVal 168
520 ATAGGCACACTAGTACGTGAGAGCTTTGGCATTAATAGCAATCAATCC 569
168 oProSerLeuMetLeuLeuLeuMetCysPhmeTrpGluThrProArg 185
570 ATGCGCTGCCCTCTCCCTTGGCTTTTATCTTCCTGAGCTCAAGAT 619
185 heLeuLeuThrGlnHisArgArgGlnGlu...AlaIleAlaLeuArgPhe 200
620 GGCTGGCAAAAGTGGGCGTGATACGAGATTTGAGAGCTGACAGTACAG 669
201 LeuTyrProIleSerGluGlnGlyTyrGluAspProProIleGlyAla 217
670 CTCGCTGGGAGAGAGCTGATATTTCCGAGAG...GCAGAGAGAT 713
217 nSerPheHisLeuAlaLeuLeuArg..... 225
714 CCAGGATTAATCGAAACTGTGGAAAGCTAACCGAAAGCCAGATGCTGG 763
226 .....GlnProGlyIleTyrLysProPheIleIleGlyValSerLeu 239
764 ATTTGTTTCAGAGAGATACATACGCTCTGTCTTATAGCTTCGCGTGG 813
240 MetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla 256
814 ATGGTGTTCAGCATTTGGAGGAATCAACGAAATATGTTTCTACACAG 863
256 uThrIlePheGlnGluAlaIleLysPheLysAspSerSerLeuAlaSerVal 273
864 CTCGATATTGAGCAGCAGAGCTTT...CCCAACAGCTTTGGATGATTA 910
273 aIValGlyValIleGlnValLeuPheThrAlaValAlaLeuIleMet 289
911 TATATCTCTCTTTCAGTGTGTAATACCTGCGCTTAATCCACGATAGTT 960
290 AspArgAlaGlyArgArgLeuLeuLeuValLeuSer.....GlyValVa 304
961 GACAGAGCGCGGAAGAAACCATTTGCTACAGGTTTTCGCAACAGGTTACT 1010
304 lMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnG 321
1011 CATAGCGCTTTGATAGCAGCGCTTCTTCTATCTCAAGGTT..... 1053
321 lGlylProGlyAsnSerSerHisValAlaIleSerAlaProValSerAla 337

```

```

1054 .....CAGCATGGCGCAGAGCATC..... 1077
338 GlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerme 354
1078 ...CCAGTC.....CTGGCTGTGTTGTTGAT 1100
354 tCysLeuPheIleAlaGlyPheAlaValGlyTyrGlyProIleProTrpL 371
1101 AATGCTGTACATAGAGATGTTTTCAGCAGAGGAATGGACCAATGGCTGG 1150
371 eLeuMetSerGluIlePheProLeuHisValLysGlyValAlaIleThrGly 387
1151 TGGTCATCTGTGAGATATTTCCCATTAACATAAAGAGTACGAGAGCC 1200
388 lLeCysValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPh 404
1201 ATGGGAGCGCTGGTAACCTGTTTGGAGCCGCGCTGTCTTCACTTT 1250
404 eSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeuAla 421
1251 CAACTTCTCATGTTC...TGGAGCTTTACGGAACCTTCTCTTATAGC 1297
421 eAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGlu 437
1298 CTCGCATCAACGCGCGCTGCGCATTTGTTCTTCATTCGCCATCGTCC 1347
438 ThrLysGlyLysThrLeuGlnGlnIleThrAla 448
1348 ACANAAGGGAAGACATTTGAGCAGACATCCAAAGCT 1380
seq_name: /SIDs2/gcdata/geneseq/geneseqn/NA2000.DAT: AAC45298
seq_documentation_block:
ID AAC45298 standard; DNA; 1685 BP.
XX
AC AAC45298;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 46007.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.

```

PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139452.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 21-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140685.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144864.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145089.

PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 28-JUL-1999; 9905-0145951.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0158293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.


```

PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

alignment_scores:
  Quality: 629.00      Length: 471
  Ratio: 2.097         Gaps: 13
  Percent Similarity: 63.694  Percent Identity: 35.456

```

alignment_block:

US-09-516-493-7 x AAC45290 ..

Align seg 1/1 to: AAC45290 from: 1 to: 1743

```

3 ValPheLeuAlaIaPheAlaIaLeuGlyProLeuSerPheGlyIph 19
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
284 GTTATATCTAGCACATCTGTCGGGTGCTTCTTTGGCGTGGTTC 333
19 eAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaIaP 36
  ::|:|||||:|||||:|||||:|||||:|||||:|||||:|||||
334 TTGTGGGGATCTCTTCACCTGCT.....CAGGCTGCATTA 371
36 roProAlaProArgLeuAspAspAlaIaIaSerTrpPheGlyAlaVal 52
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

372 GSAATGACCTTTCATTCATATAGCTGATTTTCACTCTTGGTCTTTA 421
53 ValThrLeuGlyAlaIaIaIaGlyValLeuGlyGlyTrpLeuValAs 69
  ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
422 CTAACTTTTGGCCCAATGATCGGTGCTATACACAGGGCCCTATAGCTGA 471
69 parGaIaGlyArgLysLeuSerLeuLeuLeuLeuLeuLeuValProPheValA 86
  |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
472 TTTAGTTGGACAAAGGGGCCATGACAGTTTCCTGCTGATTTTGTAG 521
86 IagGlyPheAlaValIleThrAlaIaGlnAspValITrPheLeuLeuGly 102
  ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
522 TCGGGTGGCTACCAATCATCTTGGCAAGGGAGTGGCTGTGACCTT 571
103 GlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuAlaIaPr 119
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
572 GGAAGACTGGCAAGGGATATGGAATGGAGCATTTTCTATGTGGTGC 621
119 oValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlys 136
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
622 AATCTTTATAGCAGAAATTGCACCTAAACTTTCAGAGGGGCTCTAACCA 671
136 eCys.Val..GlnLeuMetValValValGlyIleLeuLeuAlaTyrIle 151
  ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
672 CACTGAACGACAGAGATTTCTGATCTGACCTGAGAGTGGTGTCTCCAT 721
151 uAlaGlyTrpValLeuGluTrpArgTrpLeuAlaValLeuGlyCysValP 168
  :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
722 CATAGGACACTAGTGCAGTGGAGAGTCTTGGCATTAATAGGAATCATCC 771
168 roProSerLeuMetLeuLeuLeuMetCysPheMetProGluTrpProArg 184
  ||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
772 CATGGGCTGCCCTCCCTTCCTTGGCTCTTTTATCCCTGAGCTCCAGA 821
185 PheLeuLeuThrGlnHisArgArgGlnGlu..AlaIleAlaLeuArgPh 200
  ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
822 TGGCTGGCAAAAGTGGGGGTGATACGGAGTTTGAAGCTGACAGCAAA 871
200 eleuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGlu 217
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
872 GCTCCGTGGGAAGAGCGTGATTTTCGAGAG.....GCAGCAGAGA 915
217 InSerPheHisLeuAlaLeuLeuArg..... 225
  ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
916 TCCAGGATTATATCAAACTCTGAAAGGCTACCCAAAGCCAAAGTGTG 965
226 .....GlnProGlyIleTyrIlyProPheIleIleGlyValSerIle 239
  ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
966 GATTTGTTTCAGAGAGATACATACGCTCTGTTCTTATAGCTTTGGGTT 1015
239 uMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrIaG 256
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1016 CATGGCTTTCACGACTTTGGAGGAATCAGCGAATATGTTTCTACACAA 1065
256 IuThrIlePheGlnGluAlaLysPheLysAspSerSerLeuAlaSerVal 272
  ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1066 GCTCATATTTCAGCAAGAGCTTTT...CCCAAGACTTTCGATGATA 1112
273 ValValGlyValIleGlnValLeuPheThrAlaValAlaIaLeuIleMe 289
  ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1113 ATATATGCTTCTTCAGGTGATACATCACTGCGCTTATATGACCCGATGT 1162
289 IAspArgIaGlyArgArgLeuLeuValLeuSer.....GlyVal 304
  ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1163 TCACAGAGCCGGAAGAAACCTTCTACTGCTTTCTGCACAGGGTTAG 1212
304 AlMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLeuThrGln 320
  |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1213 TGATAGGCGTGTGATAGAGGCGTTCTTCTATCTCAAGCTT..... 1256
321 GlyIlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAl 337
  ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1257 .....CAGGACATGGCGCAGACGACGTC..... 1280

```


337 acInProValaspAlaservAlGlyLeuAlaTrpLeuAlaValGlySerM 354
1281CCACTC.....CTGGCTGTGTGGTA 1302
354 etCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrp 370
1303 TAATGCTTACATAGCATGCTTTCAGCAGGAATGCGCAATGCCCTGCG 1352
371 LeuLeuMetSerGluIlePheProLeuIleValGlyValAlaIleThrCl 387
1353 GTGGTCATGCTGCAGATATTTCCATAACATAAAGAGTAGAGAGAGG 1402
387 yIleCysValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluP 404
1403 CATGGCGACCGCTGTGTAAGTGTGGAGCGTGGCTGTCTTACACTT 1452
404 heSerLeuMetGluValLeuAlaGProTrpGlyAlaPheTrpLeuAla 420
1453 TCACCTTCTCATGCTGC...TCGACCTCTTACGGAACCTTCCATTTAC 1499
421 SerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProcl 437
1500 GCTGCCATCAACGCCCTGCCATTTCTTTCGTCATTGCCATGCTGCTGA 1549
437 uThrLysGlyThrLeuGluGlnIleThrAla..... 448
1550 GACAAAGGAAAGACATTGGACACATCCACACTTACTGATCCATCCATACC 1599
449HisPhe 450
1600 TCAACACTTC 1610
seq_name: /SIND2/gcgdata/geneseq/geneseq/NA2000.DAT:AAC36954
seq_documentation_block:
ID AAC36954 standard: DNA: 2134 BP.
XX
AC AAC36954:
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15657.
XX
KM Hybridisation assay: genetic mapping; gene expression control;
KW protein identification: signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000: 2000EP-0301439.
XX
PR 25-FEB-1999: 99US-0121825.
PR 05-MAR-1999: 99US-0123180.
PR 09-MAR-1999: 99US-0123548.
PR 23-MAR-1999: 99US-0125788.
PR 29-MAR-1999: 99US-0126264.
PR 01-APR-1999: 99US-0126785.
PR 06-APR-1999: 99US-0128234.
PR 08-APR-1999: 99US-0128714.
PR 16-APR-1999: 99US-0129845.
PR 19-APR-1999: 99US-0130077.
PR 21-APR-1999: 99US-0130449.
PR 23-APR-1999: 99US-0130510.
PR 28-APR-1999: 99US-0130891.
PR 30-APR-1999: 99US-0132048.
PR 30-APR-1999: 99US-0132407.

PR 04-MAY-1999: 99US-0132484.
PR 05-MAY-1999: 99US-0132485.
PR 06-MAY-1999: 99US-0132486.
PR 07-MAY-1999: 99US-0132487.
PR 11-MAY-1999: 99US-0132863.
PR 14-MAY-1999: 99US-0134256.
PR 14-MAY-1999: 99US-0134219.
PR 14-MAY-1999: 99US-0134221.
PR 14-MAY-1999: 99US-0134370.
PR 18-MAY-1999: 99US-0134768.
PR 19-MAY-1999: 99US-0134941.
PR 20-MAY-1999: 99US-0135124.
PR 21-MAY-1999: 99US-0135353.
PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136392.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137528.
PR 04-JUN-1999: 99US-0137724.
PR 07-JUN-1999: 99US-0137724.
PR 08-JUN-1999: 99US-0138094.
PR 10-JUN-1999: 99US-0138540.
PR 10-JUN-1999: 99US-0138847.
PR 14-JUN-1999: 99US-0139119.
PR 16-JUN-1999: 99US-0139452.
PR 16-JUN-1999: 99US-0139453.
PR 17-JUN-1999: 99US-0139492.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139763.
PR 21-JUN-1999: 99US-0139817.
PR 22-JUN-1999: 99US-0139899.
PR 23-JUN-1999: 99US-0140353.
PR 23-JUN-1999: 99US-0140354.
PR 24-JUN-1999: 99US-0140695.
PR 28-JUN-1999: 99US-0140823.
PR 29-JUN-1999: 99US-0140991.
PR 30-JUN-1999: 99US-0141287.
PR 01-JUL-1999: 99US-0141842.
PR 01-JUL-1999: 99US-0142154.
PR 02-JUL-1999: 99US-0142390.
PR 06-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142920.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143542.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144632.
PR 20-JUL-1999: 99US-0144864.
PR 21-JUL-1999: 99US-0144814.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.

```
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0147038.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
```

```
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

alignment_scores:

```
Quality: 547.00 Length: 462
Ratio: 2.004 Gaps: 9
Percent Similarity: 59.091 Percent Identity: 31.169
```

alignment_block:

```
US-09-516-493-7 x AAC36954 ..
```

```
Align seg 1/1 to: AAC36954 from: 1 to: 2134
```

```
10 AlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGlyTyrSerSerPr 26
   ::::::::::: :::: ::::::::::: :::::
691 GCTGCTCTGGGCTACTACTCTGGTTATCATCTCGGGGCTGTTAAAGG 740
   ::::::::::: :::::
26 CalAlleProSerLeuGlnArgAlaAlaProProAlaProArgLeuAspA 43
   ::::::::::: :::::
741 TCTCTTGATATATCTTGCTAAGATCTGGATCGCCCA.....AATA 784
   ::::::::::: :::::
43 SPAlaAlaLeuSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAla 59
   ::::::::::: :::::
785 CTGTTTCAAGAGATGATGTTGTTACTCTGCTGTCGTGTCGTCGTA 834
   ::::::::::: :::::
60 GlyGlyValLeuGlyGlyTyrLeuValAspArgAlaGlyArgLysLeu 76
   ::::::::::: :::::
835 GGTTCATTGCTACTGAGGTGTCATTAGCTGCAAAATTGGACGAACAAGAC 884
   ::::::::::: :::::
76 rLeuLeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrA 93
   ::::: :::::
885 TTTTCAATTGGATGCTATCCCTTGCCATTGAGACTTTTATTATGCAA 934
   ::::::::::: :::::
93 IaAlaGlnAspValTrpMetLeuLeuGlyArgLeuLeuThrGlyLeu 109
   ::::::::::: :::::
935 CAGCTCAGAGTGTGACACTATGATTGTGGGCGTCTGCTGCTGCAATT 984
   ::::::::::: :::::
110 AlaCysGlyValAlaSerLeuValAlaProValTyrIleSerGluIleAl 126
   ::::: :::::
985 GGAATTTGAATCTCATGACGATTTACCACTTTTACATATGCAATATTC 1034
   ::::::::::: :::::
126 aTyrProAlaValArgGlyLeuGlySerCysValIleLeuMetValV 143
   ::::::::::: :::::
1035 ACCAACTGAAATCCGTGAGGACCTGCGATCTGTGACCAAGTTGTCATCT 1084
   ::::::::::: :::::
143 aValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu..... 157
   ::::::::::: :::::
1085 GTATAGGAATACTTGCAGCTTGATAGCGATTCACCTTCGACAGAAC 1134
   ::::::::::: :::::
158 .....TPArgTrpLeuAlaValLeuGlyCysValProProSerIle 171
   ::::::::::: :::::
1135 CCTTATGATGTGAGACGACGATGTTGGTGTGAGTTATCCCTTCGTTCT 1184
   ::::::::::: :::::
171 uMetLeuLeuMetCysPheMetProGluThrProArgPheLeuLeu 188
   ::::: :::::
1185 ATTGGCCATGGAATGGCTTTTCTCCAGAAAGCCCAAGGAGGCTGCTC 1234
   ::::::::::: :::::
188 hGlnIleArgArgGlnGlnIleAla...IleAlaLeuArgPheLeuTrpGly 203
```

```
1235 AGCAGGAAAGTCTCTGAACCTGAAAGCGCATGAAACTTTGTATGCT 1284
      |||   ...   ||| ||| ||| ||| ||| ||| ||| ||| |||
204 SerGlu.....GlnG 207
      :|||
1285 AAAGAAGAGTGGTTGAACCTAGTTGCGACCTTATCAGCCTCGGCCAAGG 1334
      :|||
207 yTrpGluAspProProlIleGlyAlaGlnSerPheHisLeuAlaLeuL 224
      |   :||| |||   :|||
1335 TTCTTCGAGCGCGGAGGAGGATGCTTGATCTATTC..... 1371
      |||
224 euArgGlnProGlyLLeuTyrlsProPheLleIleGlyValSerLeuMet 240
      :|||
1372 ....ACAGCGCGCTACTCGAAAGTTGTAACCGTAGGTGCGGCTCTCTTC 1416
      :|||
241 AlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrlaGluTh 257
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1417 TTGTTTCACAGCTTAGCGCGGATTAACCCACTTGCTGATTACTCCACATC 1466
      :|||
257 rLlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValV 274
      :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1467 GGTATTCGCTAGTCCGGGAATCCATCAGATGTTCCAGCCAGTCTCTCG 1516
      :|||
274 aLGIValLleGlnValLeuPheThrAlaValAlaAlaLeuLleMetasp 290
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1517 TTGGAGCATCAAAATGCTTGGCAGCTCTTGCTTCATCGTTGATGAT 1566
      :|||
291 ArgAlaGlyArgArgLeuLeuValLeuSerGlyValValMetValPh 307
      :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1567 AAATGGGAAGCAAAAGCTTTTA..... 1590
      :|||
307 eSerThSerAlaPheGlyAlaTyrlsPheLysLeuThrgInGlyProg 324
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1591 .....CTGACAGCTTTGGTGGAA 1609
      :|||
324 LysnSerSerHisValAlaIleSerAlaProValSerAlaGlnProVal 340
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1610 TGGCTTTGTCATGCTGTACTGCTTGCCTTGCCTTCACATCGAAGCCTCTT 1659
      :|||
341 AspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPh 357
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1660 GCTGCCATTTCTGGAACC.....CTTGCCGTTGTTGGAACCTGTTCTATA 1703
      :|||
357 eLleAlaGlyPheAlaValGlyTrpGlyProLleProTrpLeuLeuMetS 374
      :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1704 TGTCCTGTCATTTCTACTGTGCTGCGCCGCTACCGGCTCTTCTTCTTC 1753
      :|||
374 ergLulIlePheProLeuHisValLysGlyValAlaThrGlyLleCysVal 390
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1754 CAGAGATATTGCAATCCGCAATCAGAGCAAAAGCGTCTTCTCTCTC 1803
      :|||
391 LeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLe 407
      :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1804 GGCATGCACCTGATATCAAACTTTGTGTCGACCTATATCTTAAGCGGT 1853
      :|||
407 uMetGluValLeuArgProTyrlGlyAlaPheTrpLeuAlaSerAlaPhe 423
      :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1854 TTGACT.....AAATTCGAATCAGCAGTGTCTACT 1885
      :|||
424 .....CysLlePheSerValLeuPheThrLeuPheCys 434
      :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1886 TGGCTTTTGTGAGTCTGCGTCTTGGCGTCTCTACATTTGCAAGAAAC 1935
      :|||
435 ValProGluThrLysGlyLysThrLeuGlnIle 446
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1936 GTCTCGAGACTAAAGTGCATCTGAGGGAATA 1971
```



```

834 GAGCTCAACCACTGGCCATGTATCGCATTTGATGCGCCAGCTGC 883
152 laGlyTrpValLeuGlu.....TrrArgTrp 160
884 TGGGC.....TTGGAGTCCCTCTGGGCACTCCAGCTGGCCACTG 927
161 LeuAlaValLeuGlyCysValProPserLeuMetLeuLeuMetCys 177
928 CTCCTGGGCTTCACAGTGTACTGCGCTTCGACCTGCTCTGCTGCC 977
177 spHeMetProGluThrProArgPheLeu...LeuThrGlnHisArgArg 193
978 CTTCTGTCCCGAGAGCCCGCTACCTTCATCATCCAGAAATTCGAGG 1027
193 lngluAla...lleAlaLeuArgPheLeu.....TrpIysSerGluGln 206
1028 GGCCTGGCCAGAAAGACTGGAAGCCGCTACAGCTGGGCGCATGTTCT 1077
207 Gly.....TrpGluAspProIleGlyAlaGluGlnSerPh 219
1078 GGAGTGTGGCTGAGCTGAGATGAGAAAGCGAGCTGGAGCTGAGCGG 1127
219 eHisLeuAlaLeuArgGlnProGlyIle.....TyrIysProP 233
1128 GCGACTGTCCCTGCTCCAGCTCTGGGCAAGCCGATACCCAGCGGACCC 1177
233 heileIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValasn 249
1178 TGATCATGTCGCTGCTGCTGACAGTGCAGCCAGCCAGCTCTGCGCATCAT 1227
250 AlaValMetPheIleGlyAlaGluThrIlePheGluGlnAlaIysPheIysAs 266
1228 GCTGTTTCTATTTATTCACACGATCTTCGACAGCAGCAGGAGTAAAGCCA 1277
266 pSerSerLeuAlaSerValValGlyValIleGlnValLeuPheThr 283
1278 GCGCTGCTATCCACACATAGAGCTGCTGCTGTAACACAGCTTCACCT 1327
283 laValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuVal 299
1328 TGGCTGCTGCTGCTGCTGCTGAGGCGGCGGCGGCGGCGGCGGCGGCTC 1377
300 leuSerGlyValValMetValPheSerThrSerAlaPheGluAlaIleYrPh 316
1378 CTG..... 1380
316 elysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSer 333
1381 .....GGCTGGGCGGCACTGTGGCTGGCCATGCTCA 1414
333 laProValSerAlaGlnProValAspAlaSerValGlyLeuAlaIlePhe 349
1415 TGACTGTGGCTGCTGCTGCTGAGGAGTTCAGGCAATGAGCTAGCTC 1464
350 AlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpG 366
1465 TCCATTGTGGCCATCTTTGGCTTCGCGCATTTTGTGAGATTGGCCCTGG 1514
366 yProIleProTrpLeuLeuMetSerGluIlePheProLeuHisValLys 383
1515 CCCCATCTCTTGGTTCATGTCGCGAGGCTTCAGCCAGGAGGAGGAGGAGG 1564
383 lValAlaThrGlyIleGysValLeuThrAsnTrpLeuMetAlaPheLeu 399
1565 CCGGACCCCATGGCTGTGGCTGTTTCTCCAACTGAGCAGCAACTTCATC 1614
400 ValThrLysGluPheSerSerLeuMetGluValIleArgProTyrGlyAl 416
1615 ATTGGCATGGTTTTCAGTATGTTGGGAGGCGCATGGGCGCTAC..... 1659
416 abHeIrrPheAlaSerAlaPheCysIlePheSerValLeu..... 429

```

```

1660 .....GTCTCTCTTCTATTGCGGCTCCTCTGCTGCGCT 1693
430 .....PheThrLeuPheCysValProGluThrLysGlyLysThrLeu 443
1694 TCTTCATCTTCACCTTTTAAGATGACCTGAAGACTCGAGGCGGACGTTT 1743
444 GluGlnIleThrAlaHisPhe 450
1744 GACGAGATCTCGGCTGCTTC 1764

```

```

seq_name: /cgn2_6/prodata/2/lna/5b_COMB.seq:US-08-928-692-9
seq_documentation block:

```

```

: Sequence 9, Application US/08928692
: Patent No. 5958727
: GENERAL INFORMATION:
: APPLICANT: Brody, Howard
: APPLICANT: Yaver, Deborah S.
: APPLICANT: Lamsa, Michael
: APPLICANT: Hansen, Kim
: TITLE OF INVENTION: Methods for Modifying the Production of
: NUMBER OF SEQUENCES: 80
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 59587270 No. 5958727/disk of No. 5958727/ln America, Inc.
: STREET: 405 Lexington Avenue
: CITY: New York
: STATE: NY
: COUNTRY: USA
: ZIP: 10174
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/928,692
: FILING DATE: 12-SEPT-1997
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Lambiris, Elias J
: REGISTRATION NUMBER: 33,728
: REFERENCE/DOCKET NUMBER: 4944.200-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-867-9655
: TELEFAX: 212-878-9655
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3000 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-928-692-9

```

alignment_scores:

```

Quality: 449.50 Length: 508
Ratio: 1.572 Gaps: 18
Percent Similarity: 56.299 Percent Identity: 29.724

```

alignment_block:

```

US-09-516-493-7 x US-08-928-692-9

```

```

Align seg 1/1 to: US-08-928-692-9 from: 1 to: 3000

```

```

8 PheAlaAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeu..... 21
11 ..... 11
585 TTTATCAACAACCTTTGGGCGCACTTCATTCGATATCATTTGGATTTACA 634
22 GlyTyrSerSerProAlaIleProSerLeuGlnArg..... 33
635 CGAGCTTGCTATGCTGAGGCTTCATATCATGCGCTGACAAATATATTT 684

```


APPLICATION NUMBER: US-09/031,392
 FILING DATE: 26-FEB-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Melkielejohn, Ph.D., Anita L.
 REGISTRATION NUMBER: 35,283
 REFERENCE/DOCKET NUMBER: 07334/072001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-8906
 TELEFAX: 617/542-8906
 TELEX: 200154
 INFORMATION FOR SEQ. ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2343 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: Coding Sequence
 LOCATION: 73...1761
 US-09-031-392-1

alignment_scores:
 Quality: 378.50 Length: 491
 Ratio: 1.418 Gaps: 13
 Percent Similarity: 54.379 Percent Identity: 26.884

alignment_block:
 US-09-516-493-7 x US-09-031-392-1 ..

Align seg 1/1 to: US-09-031-392-1 from: 1 to: 2343

```

3 ValPheLeuAlaIleAlaPheAlaIleAlaLeuGlyProLeuSerPheGlyPh 19
238 CTCCTCTGCGCTCCCTCCGCGCGCCCTTCGGCTCC.....TCCTT 278
19 eAlaLeuGlyTyrSer.....SerProAlaIleP 29
279 CCTCTAGCGGCACACCTGTCGGTGAATGCCGCCACCCCGATCA 328
29 roSerLeu.....GlnArgAlaIleProAlaPro 39
329 AGCGCTTTTCATGAGTCATGCGAAGAACGATGACGTCCA..... 372
40 ArgLeuSpAspAlaIleAlaSerTrpPheGlyAlaValAlaIleLeu 56
373 ...ATAGACCCAGACACTGACTCTGCTGCTGCTGACTGTGCAT 419
56 ValAlaIleAlaGlyValLeuGlyGlyTrpLeuValAspArgAlaGly 73
420 ATTGCGCATGCGTGACTGTGGGACCTTAATTGTGAAGATCATTTGAA 469
73 rglYLeuSer.....LeuLeuLeuSerValProPheValAla 86
470 AGGTTCTTGGAGACACACTTGTGCTGGCCATTAATGCGTTGCAATT 519
87 GlyPheAlaValIleThrAla.....AlaGlnAspValTrpPhe 59
520 TCCTCTGCTGATGCGATGCGCTGCTGCTGCGAGGACGAGCCTTTGAAT 569
99 tLeuLeuGlyGlyValrLeuLeuThrGlyLeuAlaGlyValAlaSerL 116
570 GCTCATTTGCGGAGCGTTTCATCATGGGCATAGATGAGGCGTCCCTCA 619
116 euValAlaProValTyrIleSerGlnIleAlaTyrProAlaValArgGly 132
620 GTGCGTCCCATGATGATGATGATGATGATGATGATGATGATGATGAT 669
133 LeuLeuGlySerCysValGlnLeuMetValValGlyIleLeu..... 147
670 TCCTGCGGAGGAGTGCATGCTTATATGATGATGATGATGATGATGAT 719
```

```

148 .....LeuAlaTyrLeuAlaGlyTyrValLeuGluTrpArgT 160
720 GCAGCTTCTGCGCGCCCGAGCTGTGGAAGAGAGATACCTGCGCAT 769
160 rPheAlaValLeuGlyCysValProProSerLeuMetLeuLeuMet 176
770 ACCTGTTTGGAGTATGTGTCCTGCGCGCTTGTCCAGTGTGAGCTT 819
177 CysPheMetProGluThrProArgPheLeuLeuThrGlnIleArgArg 193
820 CCTTCTTCTCCGACAGCCACGCTACCTGCTTGGAG...AAGCACA 866
193 nGlnAlaIleAlaLeuArg.....PheLeuTrpGlySerGlu 206
867 CGAGCAGAGAGTGTGAAGAGCTTCCAAAGCTTGTGGTAAAGCAGAG 916
206 InGlyTrpGluAspProProIleGlyAlaGlnGlnSerPheHis..... 220
917 TTTCCCAAGAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 966
221 .....LeuAlaLeuLeuArgGlnProGlyIleTyrLy 231
967 ATCCGCGCTGCTCCGCTGCTGAGCTGTGAGAGAGCTCCCTACCTCCG 1016
231 sPhePheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGly 248
1017 GCAGGTGTCACCGTATGTCACCATGAGCTGACAGCTGCTGCGCC 1066
248 aAspAlaValaMetPheTyrAlaGluThrIlePheGlnGlnAlaLysPhe 264
1067 TCAATGCAATTTGGTTTATACCAACAGCATCTTGAAGAGCTGGAGTC 1116
265 LysAspSerSerLeuAlaSerValVal.....GlyValIleGlnVal 279
1117 CCTCGGCAAGATCCCATACGTACCTTGAGTACAGGGGCGATCGAGC 1166
279 lLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgL 296
1167 TTTGCGTGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1216
296 euLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPhe 312
1217 CCTCTCTCAT..... 1227
313 GlyAlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSerHisVa 329
1228 .....GCTGCTTGGGCTCATGGGCTCTT 1253
329 lAlaIleSerAlaProValSerAlaGlnProValAspAlaSerValGly 346
1254 CTTTGGAGCCCGACCATCAGCTGACCTGACGAGCAGCAGCCCTGGAG 1303
346 euAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAla 362
1304 TCCTTACCTGATGATGATGATGATGATGATGATGATGATGATGAT 1353
363 ValGlyTrpGlyProIleProPheLeuLeuMetSerGlnIlePhePro 379
1354 AGTGGCCAGGAGGATCCCGCTTTCATCTTGAAGTGTGATTTCCAGA 1403
379 uHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu 396
1404 ATCTGAGCGCGCGCTGCTTTCATCATGACGACACCGTCAAGCTGCT 1453
396 eAlaPheLeuValThrLysGluPheSerSerLeuMetGlnValLeuArg 412
1454 CCAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1503
413 ProTyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerVal 429
1504 ACCTTAC...TGTTCCTAGTCTTGTCAAAATTTGTATCACAGGTGCTAT 1550
429 uPheThrLeuPheCysValProGluThrLysGlyLysThrLeuGlnGln 446
```



```

172 etleuLeuLeuMetCysPheMetProGluThrProArgPheLeuLeuThr 188
    ::::::::::::::::::::
971 TCCTGCTCTATTACTGGTCATCTGACGTCTCCAGCTGGCTGATCTCC 1020
189 GlnHisArgArgGlnGlnAlaAlaLeu..... 198
    ::::::::::::::::::::
1021 CAGAAATAGAAATGCTGAAGCCATGAGATCATTAAGCACATCCGAAGAA 1070
199 ..... ArgPheLeuTrpGlySerG 205
    ::::::::::::::::::::
1071 AAATGCAAAATCTTACCCGCTCCCTTCACAGCGCTGAGACTTGAAGAG 1120
205 GlnGlnGlyTrpGluAspProProlleGlyAlaGlnGlnSerPheHisLeu 221
    ::::::::::::::::::::
1121 AAATGGCAAGAAATGAACCT.....TCATTT...CTT 1152
222 AlaLeuLeuArgGlnProGlyIleTyrIysProPheIleIleGlyVal.. 237
    ::::::::::::::::::::
1153 GACTTGGTCAGAACTCCTCAGATTAAGAAACATATGATGATGATGTA 1202
238 ..... SerLeuMetAlaPheGlnG 244
1203 CAACGTGTACGAGCTCTGCTCTACCAAGCGCTCATTCATGCACATCG 1252
244 ILeuSerGlyValAsnAlaValMet.....PheTyrAlaGlnThrIle 258
    ::::::::::::::::::::
1253 GCCTTGCAGGTGACATATATCAGCTGATTTCTTCTACTGCGCTGCTT 1302
259 PheGluGlnAlaIysPheIysAspSerSerLeuAlaSerValValGly 275
    ::::::::::::::::::::
1303 GAATTCCACCTGCTGCTTC..... 1320
275 yValIleGlnValLeuPheThrAlaValAlaLeuIleLeuAspArg 292
    ::::::::::::::::::::
1321 .....ATGATCATCTCTATTATTCGACGCCCA 1345
292 IacIArgArgLeuLeuValLeuSerGlyValValMetValPheSer 308
    ::::::::::::::::::::
1346 TCGACGCGCTTACCTTGGCTGCATCAAAAT.....ATGTTCCAGCG 1389
309 ThrSerAlaPheGlyAlaTyrPheIysLeuThrGlnGlyProGlyAs 325
    ::::::::::::::::::::
1390 GCACCTGCTGCGCTCAGTTTATA.....CCGGTGA 1424
325 nSerSerHisValAlaIleSerAlaProValSerAlaGlnProValAsp 342
1425 T..... 1425
342 IAserValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIle 358
    ::::::::::::::::::::
1426 .....CTACAAATGCTAAATAATTATATCTCATGCTGGAGCA 1464
359 AlaIysPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerG 375
    ::::::::::::::::::::
1465 ATGGGATCACAATAGGCTAT...GAGTATAGTCTGCCGTCGAATGCTGA 1511
375 uIlePheProLeuHisValIysGlyValAlaThrGlyIleCysValLeu 392
    ::::::::::::::::::::
1512 GCTTACCCCATTCATTAGGAATCTTGGCGCTCACATCTGT..... 1554
392 hrAsnTrpLeuMetAlaPheLeuValThrIysGluPheSerSerLeuMet 408
    ::::::::::::::::::::
1555 .....TCCATCATGTCT 1566
409 Glu.....ValLeuArgProTyrGlyValPhe..... 417
    ::::::::::::::::::::
1567 GACATTGGTGGATCATCATCCGCTTCTGCTGTAACCGCTCACAACAT 1616
418 TrpLeuAlaSerAlaPheCysIlePheSerValLeu..... 429
    ::::::::::::::::::::
1617 CTGGCTGACCTCCCGCTGATGATTTCGGCGTACTGGCTTGGTTGCTG 1666

```

```

430 ..PheThrLeuPheCysValProGluThrIysGlyIysThrLeuGln 445
    ::::::::::::::::::::
1667 GAGTCTGCTGCTGCTGCTTCCAGAACTAAAGGAAACCTTTCCTCTGAG 1716
seq_name: /cgn2_6/pltdata/2/ina/6A_COMB.seq:US-08-501-572-4

```

```

seq_documentation_block:
: Sequence 4, Application US/08501572
: Patent No. 6063623
: GENERAL INFORMATION:
: APPLICANT: Koepsell, Hermann
: APPLICANT: Grundeman, Dirk
: APPLICANT: Gorboulev, Valentin
: TITLE OF INVENTION: Transport of protein which effects the
: TITLE OF INVENTION: Transport of cationic xenobiotics and/or
: TITLE OF INVENTION: DNA Sequences Encoding It And Their Use.
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finegan, Henderson, Farabow, Garrett & Dunner
: STREET: 1300 I Street, N.W., Suite 700
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/501,572
: FILING DATE:
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Toohey, Kimberlin M
: REGISTRATION NUMBER: 35,391
: REFERENCE/DOCKET NUMBER: 02481.1453-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)408-4400
: TELEFAX: (202)408-4400
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1882 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-08-501-572-4

```

```

alignment_scores:
  Quality: 245.00      Length: 475
  Ratio: 1.061         Gaps: 21
  Percent Similarity: 48.632   Percent Identity: 25.474

```

```

alignment_block:
US-09-516-493-7 x US-08-501-572-4 ..

```

```

Align seg 1/1 to: US-08-501-572-4 from: 1 to: 1882

```

```

12 LeuGlyProLeuSerPheGlyPheAlaLeuGlyTyrSerSerProAla... 27
    ::::::::::::::::::::
392 TTGGGCGCTGCGAGCATGCTGGTA.....TAGCACACCTCCCGGCTC 435
28 .....IleProSerLeuGlnArgAlaAlaProProAlaProArgLeuA 42
    ::::::::::::::::::::
436 CTCATCGTCACTGAGTTTAACCTGCTGTGACAGACGCTGAAGAGTG 485
42 sPAspAlaAlaIAserTrpPheGlyAlaValValThrLeuGlyAlaAla 58
    ::::::::::::::::::::
486 AC.....CTTTTTCAGTCTCTGTGTGAGACTTGGGCTTCTTC 520
59 AlaGlyGlyValLeuGlyIleThrLeuValAspArgAlaGlyArgIle 75

```

```

1301 ATCTTTATCCCGCATGAG.....-CSGA 1323
347 atrPrrrAlaValGIserMetCysLeuPheIIeAlaGlyPheAlaValG 364
||||| |||| | ||||| ||| ::::~
1324 CTGGTTTACACTTAACCCSTGCCCTGTCTTGCGCGATGAGGGGCACCACTTG 1373
364 IYTrglYpProIlePrTrpleuMeMetSerGiUlePheProLeuNIis 380
::: |||::: ::|||:::~
1374 TGCTGCAGATGTC...TGCGTGTGAACCCSTAGCTGTACCSTACTTC 1420
381 ValLysLIyAl.....AlaThrGIyTLecCysValLeuthras 393
::::::::::~
1421 ATCGAGATCTTGCGATGATGATGTGCTGCCCTGTGTACCTGGGTGG 1470
393 nTrPrrrMetAlaPheLeuValThrLySGluPheSerSetLeuMeGIUy 410
::: |||::| |||::: ~
1471 GATCTTACCCCCCTTCATGTGTTCCAG.....CTGATGAAG 1508
410 aLeuArGrProTYrgLYAlaPheTrPLeuAlaSerAlaPheCysTIlePhe 426
|| || ||| ::|::~
1509 TT.....TGGCAAGCCCTGGCCCTCATTTTGT 1537
427 SerValLeu.....PheThrLeuPheCysValProGI 437
::::|||~
1538 GGGGTTTTGGGCTGACTGCTGGGCCAAGACTCTTCTT...CTCCAGA 1584
437 uThrLySGlyLysTrPLeuGIUGln 445
||||||| ::|| ::|
1585 GACCAAGGtGTGGCTTGCCTGAG 1609

seq_name: /cgn2_6/plodata/2/ina/6A_COMB.seq:US-09-040-444-4

seq_documentation_block:
; Sequence 4, Application US/09040444
; Patent No. 6063766
GENERAL INFORMATION:
APPLICANT: Koepsell, Hermann
APPLICANT: Grundeman, Dirk
TITLE OF INVENTION: Transport protein which effects the
TITLE OF INVENTION: Transport of cationic xenobiotics and/or pharmaceuticals
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.
STREET: 1300 I Street, N.W., Suite 700
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,444
FILING DATE: March 18, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: O'Connor, Steven P
REGISTRATION NUMBER: 41,225
REFERENCE/DOCKET NUMBER: 2481.1453-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)408-4000
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1882 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

```

US-09-040-444-4

alignment_scores:

Quality:	245.00	Length:	475
Ratio:	1.061	Gaps:	21
Percent Similarity:	48.632	Percent Identity:	25.474

alignment_block:

us-09-516-493-7 x US-09-040-444-4

Align seg 1/1 to: US-09-040-444-4 from: 1 to: 1882

```
12 LeuGlyProLeuSerPheGlyPheAlaLeuGlyTyrSerProAla.. 27
|||||  ::::|  ||::|  ||::|  ||::|  ||::|  ||::|  ||::|
392 TTGGGGCCCTGGCGAGCATGGCTGGTA.....TAGACACTCCGGGCTC 435
28 .....lleProSerLeuAlaIleAlaIleProAlaProAlaLeu 42
:::  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|
436 CTGCATCGTCACTGAGTTTAACCTGCTGTGAGACCGCTGGAAAGTGG 485
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
486 AC.....CTTTTCAGTCTCTGCTGGAACCTGGCTTCTTC 520
59 AlaGlyGlyValLeuGlyGlyTyrPheValAspArgAlaGlyArgLysLe 75
|||||  ::::|  ||::|  ||::|  ||::|  ||::|  ||::|  ||::|
521 CTGGGCTCCCTGGTGGTGGTACATTGCAAGACGTTTGGCCCTAAGCT 570
75 uSerLeuLeuLeuSer...ValProPheValAlaGlyPheAlaVal 91
|  ||::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|
571 CTGCTTTGCTGACCAACCGCTGCTGACATCTGTCCGT.....GTGC 614
91 lenthAla...AlaGlnAspValTyrPheLeuLeuGlyArgLeuLeu 106
::|  ||::|  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
615 TAACAGCGGTGGGCCCGACACTATACATCATGTTGCTTTCCGCTGCTG 664
107 ThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProValTyrIle 123
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
665 CAGGCGATGCTGACAGAGGAGCTGCTGCCATACCTGATGATGAC 714
123 rGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 140
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
715 AGAGTTTGTGCTCTGCTACAGGAGACAGCGCATTTTGTACAGA 764
140 euMetValValGlyIleLeu...LeuAlaTyrLeuAlaGlyTyrVal 155
::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
765 TGGCTTTCACAGTGGGCTAGTGGGCTTGGCGGGTGGCTTATGCCATT 814
156 LeuGluTyrArgTyrPheAlaValLeuGlyCysValProProSerIleu 172
::|  ||::|  ||::|  ||::|  ||::|  ||::|  ||::|  ||::|
815 CCAGACTGGCGGCTGCTCCAGCTAGCTGTGCTGCCCTACCTCCCTT 864
172 lLeuLeuLeuMetCysPheMetProGluTyrProArgPheLeuLeuThr 189
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
865 CCGTCTATTACTGTTTGTCCCAATCCCGCGGTGCTGTGCC 914
189 lnhIAsArgGlnGlnAlaIleAlaLeuArgPheLeuTyrGlySerGln 205
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
915 AGAAGACACACCGCGAGCTGTCAAGATTAATGACAAATTGACAGAG 964
206 GlnGly.....TyrGluAs 210
::|  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
965 AAGGGAAGCTGCTCTCTGACCTGACAGTGTGCTTGGCTTGAGAGCA 1014
210 pProPheIleGlyAlaGlnSerPheIleIleAlaLeuLeuArgGlnp 227
::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
1015 TGCGTCAGAAAGCGAAGCTCTTCTTT...GCCGACCTGTCCGACTC 1051
227 roGlyIleTyrIlePheIleIleGlyValSerLeu..... 239
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
1062 CCAACTGAGGAAGACACCGCTATCTGATGTATGTGTTCTTCTTCT 1111
```

```
240 ..MetAlaPheGlnGlnLeu.....SerGlyValAs 249
::  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
1112 GCTGTGCTGACAGGATCATCATGACAGTGGAGACCAAGGAGGCCAA 1161
249 nAlaValMet.....PheTyrAlaGluThrIlePheGluGlnAlaLysP 264
|  ::  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
1162 COTACCTGACCTTCTTATCTCTGCTGAGAAATTCGCCGGGCT 1211
264 helYAspSerSerLeuAlaSerValValGlyValIleGlnValLeu 280
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
1212 TC..... 1213
281 PheThrAlaValAlaAlaIleuIleMetAspArgAlaGlyArgLeuLe 297
::  ::  ::  ||  ||  ||  ||  ||  ||  ||  ||
1214 .....ATCATCTGCTGACCATTCAGCCGATTTGGCCGATACCC 1254
297 uLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGly 314
:::  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
1255 AATACGGCTTCGAATCTGCTG..... 1276
314 lATyrPheLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAla 330
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
1277 .....ACGGGGGACAGCTGCTCTCATG 1300
331 lIleSerAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAl 347
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
1301 ATCTTATCCCGCATGAG.....CTGCA 1323
347 aTyrPheAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaVal 364
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
1324 CTGCTGTGAAGCTTACCCGCTGCTGCGCCGATGAGGCGGACCATTTG 1373
364 lYTrPglProIleProTyrPheLeuMetSerGluIlePheProLeuHis 380
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
1374 TGCTGCAATGCTC...TGCTGTGGAACCGTGAAGCTGATCCATTC 1420
381 ValLysGlyVal.....AlaThrGlyIleCysValLeuThrAs 393
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
1421 ATCAGAACTTGGGATGATGATGCTGCTGCCCTGATGACCTGATTC 1470
393 nTyrPheMetAlaPheLeuValThrLysGluPheSerSerLeuMetGlu 410
:  ::  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
1471 GATCTTACCCCTTCATGCTGCTCAG.....CTGATGAGAG 1508
410 alleuArgProTyrGlyAlaPheTyrPheAlaSerAlaPheCysIlePhe 426
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
1509 TT.....TGCAAGCCCTGCCCTCATTTTGT 1537
427 SerValLeu.....PheThrLeuPheCysValProG 437
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
1538 GGGTTTTGGGCTGATGCTGAGGAGCATGACTTCTT...CTGCCAGA 1584
437 uThrLysGlyLysThrLeuGln 445
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
1585 GACCAAGGCTGCTGCTTGGCTGAG 1609

seq_name: /cgn2_6/prodata/2/lna/6A_COMB.seq:us-08-501-572-5
seq_documentation_block:
: Sequence 5, Application US/08501572
: Patent No. 6063623
: GENERAL INFORMATION:
: APPLICANT: Koepsell, Hermann
: APPLICANT: Grundeman, Dirk
: APPLICANT: Gorboulev, Valentin
: TITLE OF INVENTION: Transport protein which effects the
: TITLE OF INVENTION: Transport Of Cationic Xenobiotics and/or
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
: STREET: 1300 I Street, N.W., Suite 700
: CITY: Washington
```



```

; GENERAL INFORMATION:
; APPLICANT: HITZ, WILLIAM D.
; APPLICANT: YADAV, NARENDRA S.
; TITLE OF INVENTION: ACYL-ACP THIOESTERASES GENES
; TITLE OF INVENTION: AND THEIR USE IN ALTERING PLANT
; TITLE OF INVENTION: OIL COMPOSITION
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,176
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/631,264
; FILING DATE: DECEMBER 20, 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: CHRISTENBURY, LYNNE M.
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: CR-8926-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1378 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Brassica napus
; US-08-948-176-20

```

```

Query Match          5.2%; Score 54; DB 2; Length 1378;
Best Local Similarity 100.0%; Pred. No. 5.1e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tggcgagcgagctctagacttgatcccccgggctgcaggaatcggcacag 54
   ||||||||||||||||||||||||||||||||||||||||||||||||
Db 71 TGGCGGCGGCTCTAGAACTACTGATCCCGGGCTGCAGGAATTGGCAGCAG 124

```

Search completed: February 13, 2002, 20:05:31
 Job time: 18158 sec

REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 365-262P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 536 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Trichoderma reesei
STRAIN: QM9414
US-08-911-020-1

Query Match 5.2%; Score 54; DB 2; Length 536;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgag 1029
Db 483 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCTGAG 536

RESULT 13
US-08-818-112-14
Sequence 14, Application US/08818112
Patent No. 6290969
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, David C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Vegdick, Thomas S.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,112
FILING DATE: 13-MAR-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1058 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-818-112-14

Query Match 5.2%; Score 54; DB 4; Length 1058;
Best Local Similarity 100.0%; Pred. No. 5.3e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggcgccgctctagactagtgatcccccgggctgcaggaattcgacag 54
Db 12 TGGCGCCGCTCTAGACTAGTGAATCCCGGGCTGCAGGAATTGGCAGAG 65

RESULT 14
US-08-075-533-20
Sequence 20, Application US/08075533
Patent No. 5530186
GENERAL INFORMATION:
APPLICANT: Hiltz, William D.
APPLICANT: Yadav, Narendra S.
TITLE OF INVENTION: Nucleotide Sequences of Soybean Acyl-ACP
TITLE OF INVENTION: Thioesterase Genes
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. du Pont de Nemours and Company
STREET: 1007 Market Street
CITY: Wilmington
STATE: Delaware
COUNTRY: U.S.A
ZIP: 19898

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/075,533
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/631,264
FILING DATE: 20-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Morrissey, Bruce W.
REGISTRATION NUMBER: 30,663
REFERENCE/DOCKET NUMBER: CR-8926-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (302) 992-4927
TELEFAX: (302) 892-7949

TELEX: 835420
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 1378 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Brassica napus
US-08-075-533-20

Query Match 5.2%; Score 54; DB 1; Length 1378;
Best Local Similarity 100.0%; Pred. No. 5.1e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggcgccgctctagactagtgatcccccgggctgcaggaattcgacag 54
Db 71 TGGCGCCGCTCTAGACTAGTGAATCCCGGGCTGCAGGAATTGGCAGAG 124

RESULT 15
US-08-948-176-20
Sequence 20, Application US/08948176
Patent No. 5945585

```

: APPLICANT: Dillon, Davin C.
: APPLICANT: Campos-Neto, Antonio
: APPLICANT: Houghton, Raymond
: APPLICANT: Vedvick, Thomas S.
: APPLICANT: Twardzik, Daniel R.
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
: TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
: NUMBER OF SEQUENCES: 153
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SEED and BERRY LLP
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98104-7092
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/818,112
: FILING DATE: 13-MAR-1997
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Mak1, David J.
: REGISTRATION NUMBER: 31,392
: REFERENCE/DOCKET NUMBER: 210121,411C6
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ ID NO: 101:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 500 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: US-08-818-112-101
:
: Query Match          5.2%; Score 54; DB 4; Length 500;
: Best Local Similarity 100.0%; Pred. No. 5.9e-13;
: Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: Oy 1 tggcgccgctctgaactagtgatcccccggcgtgcaggaattcgacag 54
:      |||||||||||||||||||||||||||||||||||||||||||||||
: Db 487 TGGCGCCGCTCTGAACTAGTGATCCCGCGGCTGCAGGAATTGCGACAG 434
:
: RESULT 11
: US-08-341-568-1
: Sequence 1, Application US/08341568
: Patent No. 5661021
: GENERAL INFORMATION:
: APPLICANT: Buchert, Johanna
: APPLICANT: Silka-aho, Matti
: APPLICANT: Viikari, Liisa
: APPLICANT: Penttila, Merja
: APPLICANT: Saloheimo, Anu
: APPLICANT: Marjatta, Ranna
: TITLE OF INVENTION: Mannanase enzymes, genes coding for them,
: TITLE OF INVENTION: methods for isolating the genes, and methods for bleaching 11g
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Birch, Stewart, Kolasch and Birch
: STREET: PO Box 747
: CITY: Falls Church
: STATE: VA
: COUNTRY: US
: ZIP: 22040-0747
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
```

```

: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/341,568
: FILING DATE: 22-NOV-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Murphy Jr, Gerald M
: REGISTRATION NUMBER: 28,977
: REFERENCE/DOCKET NUMBER: 365-262P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 205-8000
: TELEFAX: (703) 205-8050
: TELEX: 248345
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 536 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: ORIGINAL SOURCE:
: ORGANISM: Trichoderma reesei
: STRAIN: QM9414
:
: US-08-341-568-1
:
: Query Match          5.2%; Score 54; DB 1; Length 536;
: Best Local Similarity 100.0%; Pred. No. 5.8e-13;
: Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: Oy 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgag 1029
:      |||||||||||||||||||||||||||||||||||||||||||||||
: Db 483 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATCTGAG 536
:
: RESULT 12
: US-08-911-020-1
: Sequence 1, Application US/08911020
: Patent No. 5854047
: GENERAL INFORMATION:
: APPLICANT: Buchert, Johanna
: APPLICANT: Silka-aho, Matti
: APPLICANT: Viikari, Liisa
: APPLICANT: Penttila, Merja
: APPLICANT: Saloheimo, Anu
: APPLICANT: Marjatta, Ranna
: TITLE OF INVENTION: Mannanase enzymes, genes coding for them,
: TITLE OF INVENTION: methods for isolating the genes, and methods for bleaching
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Birch, Stewart, Kolasch and Birch
: STREET: PO Box 747
: CITY: Falls Church
: STATE: VA
: COUNTRY: US
: ZIP: 22040-0747
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/911,020
: FILING DATE: 13-AUG-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/341,568
: FILING DATE: 22-NOV-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Murphy Jr, Gerald M
```

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Trichoderma reesei
STRAIN: OM9414
US-08-341-568-3

Query Match
Best Local Similarity 100.0%; Pred. No. 6.3e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgag 1029
|||||
Db 236 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTCGAG 289

RESULT 8
US-08-911-020-3
Sequence 3, Application US/08911020
Patent No. 5854047

GENERAL INFORMATION:
APPLICANT: Buchert, Johanna
APPLICANT: Silka-aho, Matti
APPLICANT: Viikari, Liisa
APPLICANT: Penttila, Merja
APPLICANT: Saloheimo, Anu
APPLICANT: Marjatta, Ranna
TITLE OF INVENTION: Mannanase enzymes, genes coding for them,
TITLE OF INVENTION: methods for isolating the genes, and methods for bleaching
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch and Birch
STREET: PO Box 747
CITY: Falls Church
STATE: VA
COUNTRY: US
ZIP: 22040-0747

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,020
FILING DATE: 13-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/341,568
FILING DATE: 22-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Murphy JT, Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 365-262P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Trichoderma reesei

STRAIN: OM9414
US-08-911-020-3

Query Match
Best Local Similarity 100.0%; Pred. No. 6.3e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgag 1029
|||||
Db 236 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTCGAG 289

RESULT 9
US-08-946-026-23
Sequence 23, Application US/08946026
Patent No. 6034218

GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Dillon, Davin C.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Mitcham, Jennifer L.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
TITLE OF INVENTION: AND IMMUNODIAGNOSIS OF PROSTATE CANCER
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/946,026
FILING DATE: 07-OCT-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121,424C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-6031
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-946-026-23

Query Match
Best Local Similarity 100.0%; Pred. No. 6.1e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgag 1029
|||||
Db 322 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTCGAG 375

RESULT 10
US-08-818-112-101/c
Sequence 101, Application US/08818112
Patent No. 6290969

GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.

COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1(a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,406
FILING DATE: 26 April, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10814
FILING DATE: 29 October, 1993
CLASSIFICATION: 435
APPLICATION DATA:
APPLICATION NUMBER: USSN 07/968,971
FILING DATE: 30 October, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 88-1US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-424-406-6

Query Match 5.4%; Score 56; DB 1; Length 1601;
Best Local Similarity 100.0%; Pred. No. 8.1e-14;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tggcgccgctctagaactagtgatgcccgagctgcaggaattcgacagc 56
Db 7 TGGCGCGCGCTCTAGAACTAGTGATGCCCGCGCTGCAGGAATTGGCAGCGCT 62

RESULT 6
US-09-090-567-3
Sequence 3, Application US/09090567
Patent No. 5989549
GENERAL INFORMATION:
APPLICANT: Sullivan, Robert
APPLICANT: Brub, Bruno
APPLICANT: Lgar, Christine
APPLICANT: Gaudreault, Christian
TITLE OF INVENTION: Acrosomal Sperm Protein And
TITLE OF INVENTION: Uses Thereof
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swabey Ogilvy Renault
STREET: 1600 - 1981 McGill College
CITY: Montreal
STATE: QC
COUNTRY: Canada
ZIP: H3A 2Y3
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/090,567
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Kevin P
REGISTRATION NUMBER: 26,674
REFERENCE/DOCKET NUMBER: 13045-2"US" FC/CC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 514-845-7126
TELEFAX: 514-288-8389
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 912 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
US-09-090-567-3

Query Match 5.3%; Score 55; DB 2; Length 912;
Best Local Similarity 100.0%; Pred. No. 2.2e-13;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tggcgccgctctagaactagtgatgcccgagctgcaggaattcgacagc 55
Db 26 TGGCGCGCGCTCTAGAACTAGTGATGCCCGCGCTGCAGGAATTGGCAGCGC 80

RESULT 7
US-08-341-568-3
Sequence 3, Application US/08341568
Patent No. 5661021
GENERAL INFORMATION:
APPLICANT: Buchert, Johanna
APPLICANT: Silka-aho, Matti
APPLICANT: Viikari, Liisa
APPLICANT: Penttila, Merja
APPLICANT: Saloheimo, Anu
APPLICANT: Marjatta, Ranna
TITLE OF INVENTION: Mannanase enzymes, genes coding for them,
TITLE OF INVENTION: methods for isolating the genes, and methods for bleaching
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch and Birch
STREET: PO Box 747
CITY: Falls Church
STATE: VA
COUNTRY: US
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/341,568
FILING DATE: 22-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr, Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 365-262P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345

RESULT 3
US-08-905-445-7
Sequence 7, Application US/08905445
Patent No. 5864015
GENERAL INFORMATION:
APPLICANT: Pireundschuh, Michael
TITLE OF INVENTION: Hodgkin's Disease Associated Molecules And
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,445
FILING DATE: 04-AUG-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/668,128
FILING DATE: 21-JUNE-1996
APPLICATION NUMBER: 08/644,116
FILING DATE: 10-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/580,980
FILING DATE: 03-JANUARY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,328
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5864015man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5441
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2085 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-905-445-7

Query Match 5.5%; Score 57; DB 2; Length 2085;
Best Local Similarity 100.0%; Pred. No. 3,1e-14;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggcgccgctctagaactagtgatcccccgggctgcaggaattcggcacgactg 57
|||||
DB 16 TGGGGCGCGCTCTAGAACTAGTGATCCCCCGGGCTGCAGGAATTGGCAGACTG 72

RESULT 4
US-07-968-971A-10
Sequence 10, Application US/07968971A
Patent No. 5455167
GENERAL INFORMATION:
APPLICANT: Toni A. Voelker
APPLICANT: Huw Maelor Davies
TITLE OF INVENTION: Medium-Chain Thioesterases
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:

ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.7
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/968,971A
FILING DATE: 19921030
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/824,247
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/04332
FILING DATE: 21-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
REFERENCE/DOCKET NUMBER: CGNE 88
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-07-968-971A-10

Query Match 5.4%; Score 56; DB 1; Length 1601;
Best Local Similarity 100.0%; Pred. No. 8,1e-14;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggcgccgctctagaactagtgatcccccgggctgcaggaattcggcacgact 56
|||||
DB 7 TGGGGCGCGCTCTAGAACTAGTGATCCCCCGGGCTGCAGGAATTGGCAGACT 62

RESULT 5
US-08-424-406-6
Sequence 6, Application US/08424406
Patent No. 5667997
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
APPLICANT: Knutzon, Deborah S.
APPLICANT: Huw Maelor Davies
TITLE OF INVENTION: Medium Chain Thioesterases in Plants
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA

977 35 3.4 1080 1 US-08-225-757B-1 Sequence 1, Appli
978 35 3.4 1080 2 US-08-722-050-1 Sequence 1, Appli
979 35 3.4 1096 4 US-09-461-697-26 Sequence 26, Appli
980 35 3.4 1128 3 US-09-136-628-1 Sequence 1, Appli
981 35 3.4 1160 3 US-08-995-159-1 Sequence 1, Appli
982 35 3.4 1169 3 US-09-100-391-5 Sequence 5, Appli
983 35 3.4 1230 4 US-09-387-574-5 Sequence 11, Appli
984 35 3.4 1302 3 US-09-100-391-11 Sequence 7, Appli
985 35 3.4 1374 4 US-08-758-621-7 Sequence 7, Appli
986 35 3.4 1374 4 US-09-107-858-7 Sequence 1, Appli
987 35 3.4 1380 2 US-08-467-559B-1 Sequence 1, Appli
988 35 3.4 1480 2 US-09-018-584A-35 Sequence 35, Appli
989 35 3.4 1559 2 US-08-417-174-1 Sequence 1, Appli
990 35 3.4 1559 2 US-08-231-565A-1 Sequence 1, Appli
991 35 3.4 1559 2 US-09-007-961-1 Sequence 1, Appli
992 35 3.4 1559 4 US-09-267-439-1 Sequence 1, Appli
993 35 3.4 1646 4 US-08-991-786A-290 Sequence 290, App
994 35 3.4 1921 2 US-08-557-128-11 Sequence 11, Appli
995 35 3.4 1971 2 US-08-892-770-4 Sequence 4, Appli
996 35 3.4 2007 3 US-08-747-221B-36 Sequence 36, Appli
997 35 3.4 2007 3 US-08-747-221B-38 Sequence 38, Appli
998 35 3.4 2007 4 US-09-005-051-36 Sequence 36, Appli
999 35 3.4 2007 4 US-09-005-051-38 Sequence 38, Appli
1000 35 3.4 2017 4 US-09-436-983-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-680-395-4
Sequence 4, Application US/08680395
Patent No. 5892010
GENERAL INFORMATION:
APPLICANT: Gray, Joe W.
APPLICANT: Collins, Colin
APPLICANT: Hwang, Soo-In
APPLICANT: Godfrey, Tony
APPLICANT: Kowdel, David
APPLICANT: Rommens, Johanna
TITLE OF INVENTION: Genes from the 20q13 Amplicon and Their
TITLE OF INVENTION: Uses
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/680,395
FILING DATE: 15-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-068900US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2605 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

FEATURE:
NAME/KEY: -
LOCATION: 1..2605
OTHER INFORMATION: /note="cDNA clone cc43 of 4 kb
OTHER INFORMATION: transcript"

US-08-680-395-4

Query Match 5.6%; Score 58; DB 2; Length 2605;
Best Local Similarity 100.0%; Pred. No. 1.2e-14;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tggcgagccgtctagactagatgatacccgagctgcaggaatcgcagagctg 58
|||||
Db 55 tggcgagccgtctagactagatgatacccgagctgcaggaatcgcagagctg 112

RESULT 2
US-08-668-128B-7
Sequence 7, Application US/08668128B
Patent No. 5840568

GENERAL INFORMATION:
APPLICANT: Pfreundschuh, Michael
TITLE OF INVENTION: Hodgkin's Disease Associated Molecules And
TITLE OF INVENTION: Uses Thereof
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM

OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/668,128B
FILING DATE: 21-JUNE-1996

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/644,116
FILING DATE: 10-MAY-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/580,980
FILING DATE: 03-JANUARY-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,328
FILING DATE: 07-JUNE-1995

ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5840568man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: IUD 5441

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 838-3884
TELEFAX: (212) 838-9200

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2085 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-668-128B-7

Query Match 5.5%; Score 57; DB 2; Length 2085;
Best Local Similarity 100.0%; Pred. No. 3.1e-14;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tggcgagccgtctagactagatgatacccgagctgcaggaatcgcagagctg 57
|||||
Db 16 tggcgagccgtctagactagatgatacccgagctgcaggaatcgcagagctg 72

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 20:05:17 ; Search time 130.06 Seconds
(without alignments)
1805.764 Million cell updates/sec

Title: US-09-516-493-9

Perfect score: 1037

Sequence: 1 tggcgagcgctctagaacta.....aaaactcaggggggccc 1037

Scoring table: OLIGO.NUC

Gapop 60.0 , Gapext 60.0

Searched: 351203 seqs, 113238999 residues

Word size : 0

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Issued Patents.NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCMus.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	58	5.6	2605	2 US-08-680-395-4	Sequence 4, Appli
2	57	5.5	2085	2 US-08-668-1288-7	Sequence 7, Appli
3	57	5.5	2085	2 US-08-905-445-7	Sequence 7, Appli
4	56	5.4	1601	1 US-07-968-971A-10	Sequence 10, Appli
5	56	5.4	1601	1 US-08-424-406-6	Sequence 6, Appli
6	55	5.3	912	2 US-09-090-567-3	Sequence 3, Appli
7	54	5.2	289	1 US-08-341-568-3	Sequence 3, Appli
8	54	5.2	289	1 US-08-911-020-3	Sequence 3, Appli
9	54	5.2	375	3 US-08-946-026-23	Sequence 22, Appli
10	54	5.2	500	4 US-08-818-112-101	Sequence 101, App
11	54	5.2	536	2 US-08-341-568-1	Sequence 1, Appli
12	54	5.2	536	2 US-08-911-020-1	Sequence 1, Appli
13	54	5.2	1058	4 US-08-818-112-14	Sequence 14, Appli
14	54	5.2	1378	1 US-08-075-533-20	Sequence 20, Appli
15	54	5.2	1378	2 US-08-948-176-20	Sequence 20, Appli
16	54	5.2	1378	2 US-08-948-176-20	Sequence 20, Appli
17	54	5.2	1568	4 US-09-043-937A-1	Sequence 20, Appli
18	54	5.2	1568	4 US-09-043-937A-1	Sequence 20, Appli
19	54	5.2	1639	2 US-08-737-524B-1	Sequence 1, Appli
20	54	5.2	2040	2 US-08-533-669A-5	Sequence 5, Appli
21	54	5.2	2045	4 US-08-795-088A-1	Sequence 1, Appli
22	54	5.2	2233	1 US-08-496-631-1	Sequence 1, Appli
23	54	5.2	2503	1 US-08-472-934-3	Sequence 11, Appli
24	54	5.2	2503	1 US-08-472-934-3	Sequence 11, Appli
25	54	5.2	2503	2 US-08-323-460A-3	Sequence 3, Appli
26	54	5.2	2503	2 US-08-461-146C-3	Sequence 3, Appli
27	54	5.2	2503	2 US-08-461-146C-11	Sequence 11, Appli

28	54	5.2	2503	3 US-08-461-145C-3	Sequence 3, Appli
29	54	5.2	2503	3 US-08-461-145C-11	Sequence 11, Appli
30	54	5.2	2775	4 US-08-730-771-1	Sequence 1, Appli
31	54	5.2	2775	4 US-09-060-208-1	Sequence 1, Appli
32	54	5.2	3089	1 US-08-472-934-5	Sequence 5, Appli
33	54	5.2	3089	2 US-08-323-460A-5	Sequence 5, Appli
34	54	5.2	3089	3 US-08-461-146C-5	Sequence 5, Appli
35	54	5.2	3089	2 US-08-461-146C-5	Sequence 5, Appli
36	54	5.2	3328	4 US-08-960-048-1	Sequence 4, Appli
37	53	5.1	1223	3 US-09-154-874-4	Sequence 4, Appli
38	53	5.1	1899	1 US-08-634-331-2	Sequence 2, Appli
39	52	5.0	742	1 US-07-847-010-12	Sequence 12, Appli
40	52	5.0	974	2 US-08-504-459-13	Sequence 13, Appli
41	52	5.0	1146	4 US-08-893-654B-5	Sequence 5, Appli
42	52	5.0	2065	2 US-08-968-751-1	Sequence 1, Appli
43	51	4.9	746	4 US-09-013-810-1	Sequence 1, Appli
44	51	4.9	780	2 US-08-540-804-36	Sequence 36, Appli
45	51	4.9	780	4 US-08-590-399-36	Sequence 36, Appli
46	51	4.9	897	4 US-08-899-330-14	Sequence 14, Appli
47	51	4.9	1046	1 US-08-361-467B-4	Sequence 4, Appli
48	51	4.9	1046	1 US-08-484-332C-4	Sequence 4, Appli
49	51	4.9	1319	2 US-08-504-459-7	Sequence 7, Appli
50	51	4.9	1332	2 US-09-057-762-1	Sequence 1, Appli
51	51	4.9	1332	3 US-08-326-119A-1	Sequence 1, Appli
52	51	4.9	1382	4 US-09-370-253-1	Sequence 1, Appli
53	51	4.9	2114	1 US-07-803-622E-6	Sequence 6, Appli
54	51	4.9	2407	4 US-09-370-807-7	Sequence 7, Appli
55	51	4.9	2567	3 US-08-993-260-4	Sequence 4, Appli
56	51	4.9	3441	2 US-08-742-753-1	Sequence 1, Appli
57	50	4.8	80	1 US-07-920-281C-25	Sequence 25, Appli
58	50	4.8	80	4 US-08-466-277-25	Sequence 25, Appli
59	50	4.8	90	1 US-08-677-944-1	Sequence 1, Appli
60	50	4.8	234	1 US-08-485-455D-60	Sequence 60, Appli
61	50	4.8	234	2 US-08-482-130C-60	Sequence 60, Appli
62	50	4.8	234	2 US-08-484-211C-60	Sequence 60, Appli
63	50	4.8	234	3 US-08-906-719-60	Sequence 60, Appli
64	50	4.8	234	3 US-08-906-616-60	Sequence 60, Appli
65	50	4.8	234	3 US-08-817-795-60	Sequence 60, Appli
66	50	4.8	234	3 US-08-485-443B-60	Sequence 60, Appli
67	50	4.8	234	3 US-08-639-075A-60	Sequence 60, Appli
68	50	4.8	234	4 US-09-012-431-60	Sequence 60, Appli
69	50	4.8	234	4 US-09-012-692-60	Sequence 60, Appli
70	50	4.8	234	4 US-08-906-613-60	Sequence 60, Appli
71	50	4.8	703	5 PCT-US95-14442A-60	Sequence 60, Appli
72	50	4.8	860	1 US-08-181-271A-33	Sequence 33, Appli
73	50	4.8	860	1 US-08-449-315-33	Sequence 33, Appli
74	50	4.8	860	1 US-08-444-803-33	Sequence 33, Appli
75	50	4.8	860	1 US-08-449-043-33	Sequence 33, Appli
76	50	4.8	860	1 US-08-456-265A-33	Sequence 33, Appli
77	50	4.8	860	1 US-08-456-265A-33	Sequence 33, Appli
78	50	4.8	860	1 US-08-455-416-33	Sequence 33, Appli
79	50	4.8	860	1 US-08-455-244-33	Sequence 33, Appli
80	50	4.8	860	1 US-08-454-876-33	Sequence 33, Appli
81	50	4.8	860	2 US-08-456-364-33	Sequence 33, Appli
82	50	4.8	860	2 US-08-456-262-33	Sequence 33, Appli
83	50	4.8	860	2 US-08-456-240-33	Sequence 33, Appli
84	50	4.8	860	2 US-08-455-726-33	Sequence 33, Appli
85	50	4.8	860	2 US-08-971-217-33	Sequence 33, Appli
86	50	4.8	860	4 US-09-335-600-33	Sequence 33, Appli
87	50	4.8	972	1 US-07-915-924-1	Sequence 1, Appli
88	50	4.8	972	1 US-08-325-743-1	Sequence 1, Appli
89	50	4.8	1001	1 US-08-728-259A-10	Sequence 10, Appli
90	50	4.8	1001	2 US-08-473-486-10	Sequence 10, Appli
91	50	4.8	1098	3 US-09-248-335-35	Sequence 35, Appli
92	50	4.8	1332	4 US-09-333-423-1	Sequence 1, Appli
93	50	4.8	1395	2 US-08-553-367A-1	Sequence 1, Appli
94	50	4.8	1395	4 US-09-295-306-1	Sequence 1, Appli
95	50	4.8	1441	4 US-08-821-994-63	Sequence 63, Appli
96	50	4.8	1602	1 US-08-530-950-3	Sequence 3, Appli
97	50	4.8	1602	1 US-08-888-429A-3	Sequence 3, Appli
98	50	4.8	1602	4 US-09-149-879-3	Sequence 3, Appli
99	50	4.8	1929	4 US-09-146-950-1	Sequence 1, Appli
100	50	4.8	2082	2 US-08-785-310A-2	Sequence 2, Appli

CC the new polynucleotides. Specific uses are described for each of the 186
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAV59511 for described uses).

XX
SQ Sequence 2346 Bp: 545 A: 614 C: 616 G: 566 T: 5 other:

Query Match 5.7%; Score 59; DB 19; Length 2346;
Best Local Similarity 100.0%; Pred. No. 7.5e-12;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 976 aaactcgaaggagg 1034
|||||
Db 2251 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgaaggagg 2309

Search completed: February 13, 2002, 20:10:32
Job time: 13484 sec

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 PN WC9839448-A2.
 XX
 PD 11-SEP-1998.
 XX
 PF 06-MAR-1998; 98WO-US04493.
 XX
 PR 02-OCT-1997; 97US-0061060.
 PR 07-MAR-1997; 97US-0038621.
 PR 07-MAR-1997; 97US-0040161.
 PR 07-MAR-1997; 97US-0040162.
 PR 07-MAR-1997; 97US-0040163.
 PR 07-MAR-1997; 97US-0040333.
 PR 07-MAR-1997; 97US-0040334.
 PR 07-MAR-1997; 97US-0040336.
 PR 07-MAR-1997; 97US-0040626.
 PR 11-APR-1997; 97US-0043311.
 PR 11-APR-1997; 97US-0043312.
 PR 11-APR-1997; 97US-0043313.
 PR 11-APR-1997; 97US-0043314.
 PR 11-APR-1997; 97US-0043568.
 PR 11-APR-1997; 97US-0043569.
 PR 11-APR-1997; 97US-0043576.
 PR 11-APR-1997; 97US-0043578.
 PR 11-APR-1997; 97US-0043580.
 PR 11-APR-1997; 97US-0043669.
 PR 11-APR-1997; 97US-0043670.
 PR 11-APR-1997; 97US-0043671.
 PR 11-APR-1997; 97US-0043672.
 PR 11-APR-1997; 97US-0043674.
 PR 23-MAY-1997; 97US-0047492.
 PR 23-MAY-1997; 97US-0047500.
 PR 23-MAY-1997; 97US-0047501.
 PR 23-MAY-1997; 97US-0047502.
 PR 23-MAY-1997; 97US-0047503.
 PR 23-MAY-1997; 97US-0047581.
 PR 23-MAY-1997; 97US-0047582.
 PR 23-MAY-1997; 97US-0047583.
 PR 23-MAY-1997; 97US-0047584.
 PR 23-MAY-1997; 97US-0047585.
 PR 23-MAY-1997; 97US-0047586.
 PR 23-MAY-1997; 97US-0047587.
 PR 23-MAY-1997; 97US-0047588.
 PR 23-MAY-1997; 97US-0047589.
 PR 23-MAY-1997; 97US-0047590.
 PR 23-MAY-1997; 97US-0047592.
 PR 23-MAY-1997; 97US-0047593.
 PR 23-MAY-1997; 97US-0047594.
 PR 23-MAY-1997; 97US-0047595.
 PR 23-MAY-1997; 97US-0047596.
 PR 23-MAY-1997; 97US-0047597.
 PR 23-MAY-1997; 97US-0047598.
 PR 23-MAY-1997; 97US-0047599.
 PR 23-MAY-1997; 97US-0047600.
 PR 23-MAY-1997; 97US-0047601.
 PR 23-MAY-1997; 97US-0047612.
 PR 23-MAY-1997; 97US-0047613.
 PR 23-MAY-1997; 97US-0047614.
 PR 23-MAY-1997; 97US-0047615.
 PR 23-MAY-1997; 97US-0047617.
 PR 23-MAY-1997; 97US-0047618.
 PR 23-MAY-1997; 97US-0047632.

PR 23-MAY-1997; 97US-0047633.
 PR 06-JUN-1997; 97US-0048964.
 PR 06-JUN-1997; 97US-0048974.
 PR 13-JUN-1997; 97US-0049610.
 PR 08-JUL-1997; 97US-0051926.
 PR 16-JUL-1997; 97US-0052874.
 PR 18-AUG-1997; 97US-0053724.
 PR 22-AUG-1997; 97US-0056630.
 PR 22-AUG-1997; 97US-0056631.
 PR 22-AUG-1997; 97US-0056632.
 PR 22-AUG-1997; 97US-0056636.
 PR 22-AUG-1997; 97US-0056637.
 PR 22-AUG-1997; 97US-0056662.
 PR 22-AUG-1997; 97US-0056664.
 PR 22-AUG-1997; 97US-0056845.
 PR 22-AUG-1997; 97US-0056862.
 PR 22-AUG-1997; 97US-0056864.
 PR 22-AUG-1997; 97US-0056872.
 PR 22-AUG-1997; 97US-0056874.
 PR 22-AUG-1997; 97US-0056875.
 PR 22-AUG-1997; 97US-0056876.
 PR 22-AUG-1997; 97US-0056877.
 PR 22-AUG-1997; 97US-0056878.
 PR 22-AUG-1997; 97US-0056879.
 PR 22-AUG-1997; 97US-0056880.
 PR 22-AUG-1997; 97US-0056881.
 PR 22-AUG-1997; 97US-0056882.
 PR 22-AUG-1997; 97US-0056884.
 PR 22-AUG-1997; 97US-0056886.
 PR 22-AUG-1997; 97US-0056887.
 PR 22-AUG-1997; 97US-0056888.
 PR 22-AUG-1997; 97US-0056889.
 PR 22-AUG-1997; 97US-0056892.
 PR 22-AUG-1997; 97US-0056893.
 PR 22-AUG-1997; 97US-0056894.
 PR 22-AUG-1997; 97US-0056903.
 PR 22-AUG-1997; 97US-0056908.
 PR 22-AUG-1997; 97US-0056909.
 PR 22-AUG-1997; 97US-0056910.
 PR 22-AUG-1997; 97US-0056911.
 PR 05-SEP-1997; 97US-0057650.
 PR 05-SEP-1997; 97US-0057659.
 PR 05-SEP-1997; 97US-0057761.
 PR 12-SEP-1997; 97US-0058785.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Bednarek DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA,
 PI Feng P, Ferrite AM, Fischer CL, Florence KA, Greene JM, Hu JS,
 PI Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA,
 PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
 XX
 DR WPI; 1998-506364/43.
 DR P-PSDB; AAW74910.
 XX
 PT New isolated human genes and the secreted polypeptide(s) they encode
 PT - useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 XX
 PS Claim 1; Page 423-424; 721pp; English.
 XX
 CC This sequence represents a nucleic acid molecule designated Gene 183
 CC from the human cDNA clone HTOAF35 (deposited as clone ATCC 97904 and
 CC ATCC 209050) which encodes a secreted human protein. The gene can be
 CC used to generate fusion proteins by linking to the gene to a human
 CC immunoglobulin Fc portion (e.g. AAY5502) for increasing the stability of
 CC the fused protein as compared to the human protein only.
 CC The invention relates to 186 novel genes and their fragments (nucleic
 CC acid sequences: AAY59511-V59812; amino acid sequences AAW74731-W75026)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in


```
XX 01-JUN-2001 (first entry)
DE Human secreted protein cDNA, SEQ ID NO: 45.
XX
XX
XX Human: secreted protein; immunomodulatory; antisclerotic;
XX dermatological; antiinflammatory; anti-HIV; cytostatic; cardiant;
XX vascular; anti-angiogenic; ophthalmological; neuroprotectant;
XX nootropic; anticonvulsant; antiAlzheimer's; antiparkinsonian;
XX antimicrobial; vulnerary; vaccine; gene therapy; cancer;
XX protein coordinate data; infection; ss.
XX
XX Homo sapiens.
XX
XX WO200121658-A1.
XX
XX 29-MAR-2001.
XX
XX 22-SEP-2000; 2000WO-US26013.
XX
XX 24-SEP-1999; 99US-0155709.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;
XX Lafleur DM, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;
XX Young PE, Wei P, Florence KA;
XX
XX WPI: 2001-235311/24.
XX
XX Nucleic acids encoding 32 human secreted polypeptides, useful for
XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's
XX disease and diabetic retinopathy -
XX
XX
XX Claim 1: Page 743; 890pp; English.
XX
XX The present sequence encodes one of 32 novel human secreted polypeptides.
XX The nucleic acid molecules and polypeptides they encode may be used in
XX the prevention, diagnosis and treatment of diseases such as
XX immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus
XX and human immuno-deficiency virus (HIV) infections), hyperproliferative
XX disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases
XX (e.g. Schmitz syndrome, Chaga's cardiomyopathy and coronary
XX arteriosclerosis), angiogenic disorders (e.g. corneal graft
XX neovascularisation and diabetic retinopathy), neurological disorders
XX (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
XX infectious diseases and/or for promoting wound healing, regeneration
XX and/or chemotaxis. The nucleic acid molecules may be used to produce the
XX secreted polypeptides. They may also be used as DNA probes in diagnostic
XX assays to detect and quantitate the presence of similar nucleic acid
XX sequences in samples. The polypeptides may be used as antigens in the
XX production of antibodies and in assays to identify modulators of
XX their expression and activity.
XX
XX Sequence 1071 BP; 256 A; 310 C; 321 G; 180 T; 4 other;
XX
XX
XX Query Match 5.7%; Score 59; DB 22; Length 1071;
XX Best Local Similarity 100.0%; Pred. No. 9.1e-12;
XX Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 979 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgaggggggcc 1037
XX |
XX 987 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgaggggggcc 1045
XX
XX
XX RESULT 14
XX AAF97920
XX ID AAF97920 standard; cDNA; 1149 BP.
XX
XX AC AAF97920;
XX
XX 01-JUN-2001 (first entry)
```

```
XX Human secreted protein cDNA, SEQ ID NO: 47.
DE
XX
XX Human: secreted protein; immunomodulatory; antisclerotic;
XX dermatological; antiinflammatory; anti-HIV; cytostatic; cardiant;
XX vascular; anti-angiogenic; ophthalmological; neuroprotectant;
XX nootropic; anticonvulsant; antiAlzheimer's; antiparkinsonian;
XX antimicrobial; vulnerary; vaccine; gene therapy; cancer;
XX protein coordinate data; infection; ss.
XX
XX Homo sapiens.
XX
XX WO200121658-A1.
XX
XX 29-MAR-2001.
XX
XX 22-SEP-2000; 2000WO-US26013.
XX
XX 24-SEP-1999; 99US-0155709.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;
XX Lafleur DM, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;
XX Young PE, Wei P, Florence KA;
XX
XX WPI: 2001-235311/24.
XX
XX Nucleic acids encoding 32 human secreted polypeptides, useful for
XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's
XX disease and diabetic retinopathy -
XX
XX
XX Claim 1: Page 744; 890pp; English.
XX
XX The present sequence encodes one of 32 novel human secreted polypeptides.
XX The nucleic acid molecules and polypeptides they encode may be used in
XX the prevention, diagnosis and treatment of diseases such as
XX immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus
XX and human immuno-deficiency virus (HIV) infections), hyperproliferative
XX disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases
XX (e.g. Schmitz syndrome, Chaga's cardiomyopathy and coronary
XX arteriosclerosis), angiogenic disorders (e.g. corneal graft
XX neovascularisation and diabetic retinopathy), neurological disorders
XX (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
XX infectious diseases and/or for promoting wound healing, regeneration
XX and/or chemotaxis. The nucleic acid molecules may be used to produce the
XX secreted polypeptides. They may also be used as DNA probes in diagnostic
XX assays to detect and quantitate the presence of similar nucleic acid
XX sequences in samples. The polypeptides may be used as antigens in the
XX production of antibodies and in assays to identify modulators of
XX their expression and activity.
XX
XX Sequence 1149 BP; 271 A; 334 C; 352 G; 191 T; 1 other;
XX
XX
XX Query Match 5.7%; Score 59; DB 22; Length 1149;
XX Best Local Similarity 100.0%; Pred. No. 9e-12;
XX Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 979 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgaggggggcc 1037
XX |
XX 1065 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgaggggggcc 1123
XX
XX
XX RESULT 15
XX AAV59693
XX ID AAV59693 standard; DNA; 2346 BP.
XX
XX AC AAV59693;
XX
XX 19-JAN-1999 (first entry)
XX
XX Human secreted protein gene 183 clone HTOAF35.
```

```

PR 07-MAR-1997; 97US-0040162.
PR 07-MAR-1997; 97US-0040163.
PR 07-MAR-1997; 97US-0040333.
PR 07-MAR-1997; 97US-0040334.
PR 07-MAR-1997; 97US-0040336.
PR 07-MAR-1997; 97US-0040626.
PR 11-APR-1997; 97US-0043311.
PR 11-APR-1997; 97US-0043312.
PR 11-APR-1997; 97US-0043313.
PR 11-APR-1997; 97US-0043314.
PR 11-APR-1997; 97US-0043568.
PR 11-APR-1997; 97US-0043569.
PR 11-APR-1997; 97US-0043576.
PR 11-APR-1997; 97US-0043578.
PR 11-APR-1997; 97US-0043580.
PR 11-APR-1997; 97US-0043669.
PR 11-APR-1997; 97US-0043670.
PR 11-APR-1997; 97US-0043671.
PR 11-APR-1997; 97US-0043672.
PR 11-APR-1997; 97US-0043674.
PR 23-MAY-1997; 97US-0047492.
PR 23-MAY-1997; 97US-0047500.
PR 23-MAY-1997; 97US-0047501.
PR 23-MAY-1997; 97US-0047502.
PR 23-MAY-1997; 97US-0047503.
PR 23-MAY-1997; 97US-0047581.
PR 23-MAY-1997; 97US-0047582.
PR 23-MAY-1997; 97US-0047583.
PR 23-MAY-1997; 97US-0047584.
PR 23-MAY-1997; 97US-0047585.
PR 23-MAY-1997; 97US-0047586.
PR 23-MAY-1997; 97US-0047587.
PR 23-MAY-1997; 97US-0047588.
PR 23-MAY-1997; 97US-0047589.
PR 23-MAY-1997; 97US-0047590.
PR 23-MAY-1997; 97US-0047592.
PR 23-MAY-1997; 97US-0047593.
PR 23-MAY-1997; 97US-0047594.
PR 23-MAY-1997; 97US-0047595.
PR 23-MAY-1997; 97US-0047596.
PR 23-MAY-1997; 97US-0047597.
PR 23-MAY-1997; 97US-0047598.
PR 23-MAY-1997; 97US-0047599.
PR 23-MAY-1997; 97US-0047600.
PR 23-MAY-1997; 97US-0047601.
PR 23-MAY-1997; 97US-0047612.
PR 23-MAY-1997; 97US-0047613.
PR 23-MAY-1997; 97US-0047614.
PR 23-MAY-1997; 97US-0047615.
PR 23-MAY-1997; 97US-0047617.
PR 23-MAY-1997; 97US-0047618.
PR 23-MAY-1997; 97US-0047632.
PR 23-MAY-1997; 97US-0047633.
PR 06-JUN-1997; 97US-0048974.
PR 06-JUN-1997; 97US-0048974.
PR 13-JUN-1997; 97US-0049610.
PR 08-JUL-1997; 97US-0051926.
PR 16-JUL-1997; 97US-0052874.
PR 16-AUG-1997; 97US-0055724.
PR 22-AUG-1997; 97US-0056630.
PR 22-AUG-1997; 97US-0056631.
PR 22-AUG-1997; 97US-0056632.
PR 22-AUG-1997; 97US-0056636.
PR 22-AUG-1997; 97US-0056637.
PR 22-AUG-1997; 97US-0056662.
PR 22-AUG-1997; 97US-0056664.
PR 22-AUG-1997; 97US-0056665.
PR 22-AUG-1997; 97US-0056666.
PR 22-AUG-1997; 97US-0056667.
PR 22-AUG-1997; 97US-0056672.
PR 22-AUG-1997; 97US-0056874.
PR 22-AUG-1997; 97US-0056875.
PR 22-AUG-1997; 97US-0056876.

PR 22-AUG-1997; 97US-0056877.
PR 22-AUG-1997; 97US-0056878.
PR 22-AUG-1997; 97US-0056879.
PR 22-AUG-1997; 97US-0056880.
PR 22-AUG-1997; 97US-0056881.
PR 22-AUG-1997; 97US-0056882.
PR 22-AUG-1997; 97US-0056884.
PR 22-AUG-1997; 97US-0056886.
PR 22-AUG-1997; 97US-0056887.
PR 22-AUG-1997; 97US-0056888.
PR 22-AUG-1997; 97US-0056889.
PR 22-AUG-1997; 97US-0056892.
PR 22-AUG-1997; 97US-0056893.
PR 22-AUG-1997; 97US-0056894.
PR 22-AUG-1997; 97US-0056903.
PR 22-AUG-1997; 97US-0056908.
PR 22-AUG-1997; 97US-0056909.
PR 22-AUG-1997; 97US-0056910.
PR 22-AUG-1997; 97US-0056911.
PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057659.
PR 05-SEP-1997; 97US-0057761.
PR 12-SEP-1997; 97US-0058785.

PR (HUMA-) HUMAN GENOME SCI INC.
PR PA
PR XX
PR Bednarik DP, Brewer IA, Carter KC, Duan R, Ebner R, Endress GA;
PR PI Feng P, Ferrie AM, Florence KA, Greene JM, Hu JS;
PR PI Kyaw H, Lafleur DM, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
PR PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
PR DR WPI: 1998-506364/43.
PR DR P-PSDB: AAW75023.
PR XX
PR PT New isolated human genes and the secreted polypeptide(s) they encode
PR PT disorders, immune diseases, inflammation or blood disorders
PR PT disorders, immune diseases, inflammation or blood disorders
PR XX
PR XX Claim 1: Page 540-541; 721pp; English.
PR CC This sequence represents a nucleic acid molecule designated Gene 183
PR CC from the human cDNA clone HMOAF35 (deposited as clone ATCC 97904 and
PR CC ATCC 209050) which encodes a secreted human protein. The gene can be
PR CC used to generate fusion proteins by linking to the gene to a human
PR CC immunoglobulin Fc portion (e.g. AAV59502) for increasing the stability of
PR CC the fused protein as compared to the human protein only.
PR CC The invention relates to 186 novel genes and their fragments (nucleic
PR CC acid sequences: AAV59511-V59812; amino acid sequences AAW74731-W75026)
PR CC which are useful for preventing, treating or ameliorating medical
PR CC conditions e.g. by protein or gene therapy. Also, pathological
PR CC conditions can be diagnosed by determining the amount of the new
PR CC polypeptides in a sample or by determining the presence of mutations in
PR CC the new polynucleotides. Specific uses are described for each of the 186
PR CC polynucleotides, based on which tissues they are most highly expressed in
PR CC (see AAV59511 for described uses).
PR XX
PR XX Sequence 2025 BP; 489 A; 527 C; 517 G; 490 T; 2 other:

Query Match 5.8%; Score 60; DB 19; Length 2025;
Best Local Similarity 100.0%; Pred. No. 3 6e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 975 caaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagggggg 1034
Db 1929 caaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagggggg 1988

RESULT 13
AAAF97918
ID AAF97918 standard; cDNA; 1071 BP.
XX
AC AAF97918;

```

PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers -
XX
PS Claim 1; Page 529; 607pp; English.
XX
CC AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted
CC protein genes, and AAB87342-AAB87413 represent the proteins they encode.
CC AAB87414-AAB87454 represent human secreted protein fragments. The genes
CC and their corresponding secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 52 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein-encoding cDNA of the invention.
XX
SQ Sequence 1086 BP; 260 A; 316 C; 326 G; 184 T; 0 other;

Query Match 5.8%; Score 60; DB 22; Length 1086;
Best Local Similarity 100.0%; Pred. No. 4.1e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0Y 978 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagggggggc 1037
|||||
Db 1001 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagggggggc 1060

RESULT 11
AAZ48807
ID AAZ48807 standard; cDNA; 1574 BP.
XX
AC AAZ48807;
XX
DT 21-MAR-2000 (first entry)
XX
DE Rice inositol 1,3,4-triphosphate 5/6-kinase coding sequence.
XX
XX Inositol 1,3,4-triphosphate 5/6-kinase; phytic acid; genetic mapping;
KM myo-inositol 1,2,3,4,5,6-hexaphosphate; biosynthetic enzyme; phytate;
KM animal feed; ss.
XX
XX Ofyza saliva.
XX
XX OS
XX PN MO9955879-A1.
XX
XX PD 04-NOV-1999.
XX
XX PF 22-APR-1999; 99WO-US08790.
XX
XX PR 24-APR-1998; 98US-0082960.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX

PI Calhoun RE, Carlson TJ, Hitz WD, Pearlstein RW;
XX
XX WPI: 2000-072179/06.
DR P-PSDB; AAV59423.
XX
XX
PS Novel phytic acid biosynthetic enzymes used to alter the level of the
PT enzyme in transgenic plants -
XX
XX Claim 2; Page 46-47; 63pp; English.
XX
CC This sequence encodes an inositol 1,3,4-triphosphate 5/6-kinase (ITK)
CC protein of the invention. ITK is a phytic acid (MYO-inositol
CC 1,2,3,4,5,6-hexaphosphate) biosynthetic enzyme. The ITK enzymes of the
CC invention may be prepared recombinantly and used to raise antibodies,
CC which are used for detecting the enzymes in situ in cells or in vitro
CC in cell extracts. The polynucleotides of the invention may be used to
CC create transgenic plants in which the ITK levels are present at higher
CC or lower levels than normal, or in cell types or developmental processes
CC where they are not normally found. This would alter the level of
CC 1,3,4-triphosphate 5/6 kinase found in those cells. In addition, it may
CC be desirable to eliminate expression of ITK genes for certain
CC applications. The polynucleotides also serve as a source of probes and
CC primers, which are useful for genetic mapping, as markers for traits
CC linked to those genes, and to isolate homologous sequences from other
CC species, as well as for physical mapping. Decreased amounts of phytate
CC and increased amounts of available phosphate in animal feeds such as corn
CC would lead to improved feed efficiency. The proteins of the present
CC invention lead to a better understanding of the phytic acid biosynthesis
CC pathway, allowing it to be exploited for commercial uses, e.g. in animal
CC feeds.
XX
SQ Sequence 1574 BP; 370 A; 551 C; 382 G; 271 T; 0 other;

Query Match 5.8%; Score 60; DB 21; Length 1574;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0Y 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagggggggg 1035
|||||
Db 1502 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagggggggg 1561

RESULT 12
AAV59809
ID AAV59809 standard; DNA; 2025 BP.
XX
AC AAV59809;
XX
DT 19-JAN-1999 (first entry)
XX
DE Human secreted protein gene 183 clone HTOAR35.
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KM diagnosis; tissue; cancer; neurodegenerative disorder; leukaemia;
KM developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KM immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KM inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KM cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KM osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KM endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
XX Homo sapiens.
XX
XX OS
XX PN MO9839448-A2.
XX
XX PD 11-SEP-1998.
XX
XX PF 06-MAR-1998; 98WO-US04493.
XX
XX PR 02-OCT-1997; 97US-0061060.
XX
XX PR 07-MAR-1997; 97US-0038621.
XX
XX PR 07-MAR-1997; 97US-0040161.

PT disorders, immune diseases, inflammation or blood disorders
 XX
 PS Claim 1; Page 271; 447pp; English.
 CC This sequence represents a nucleic acid molecule which encodes a
 CC secreted human protein. The gene number, and the clone it is derived
 CC from, are detailed in the descriptor line. The gene can be used to
 CC generate fusion proteins by linking to the gene to a human immunoglobulin
 CC Fc portion (e.g. AAV34145) for increasing the stability of the fused
 CC protein as compared to the human protein only.
 CC The invention relates to 70 novel genes and their fragments (nucleic
 CC acid sequences: AAV34154-V34276; amino acid sequences AAW75057-W75179)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 70
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAV34154 for described uses).
 CC
 XX
 SQ Sequence 737 BP; 160 A; 227 C; 228 G; 109 T; 13 other;

Query Match 5.8%; Score 60; DB 19; Length 737;
 Best Local Similarity 100.0%; Pred. No. 4.5e-12;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 978 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacctcgaggggggccc 1037
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 644 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacctcgaggggggccc 703

RESULT 9
 AAF97921
 ID AAF97921 standard; cDNA; 1086 BP.
 XX
 AC AAF97921;
 XX
 DT 01-JUN-2001 (first entry)
 DE Human secreted protein cDNA, SEQ ID NO: 48.
 DE
 XX Human: secreted protein; immunomodulatory; antisclerotic;
 KW dermatological; antinflammatory; anti-HIV; cytostatic; cardiact;
 KW vascular; anti-angiogenic; ophthalmological; neuroprotectant;
 KW neurotropic; anticonvulsant; antialzheimers; antiparkinsonian;
 KW antimicrobial; vulnery; vaccine; gene therapy; cancer;
 KW protein coordinate data; infection; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200121658-A1.
 PD 29-MAR-2001.
 PD
 XX 22-SEP-2000; 2000WO-US26013.
 PF
 XX 24-SEP-1999; 99US-0155709.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX NI J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;
 PI Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;
 PI Young PE, Wei P, Florence KA;
 XX
 XX WPI; 2001-235311/24.

PT Nucleic acids encoding 32 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy -
 XX
 XX Claim 1; Page 744; 890pp; English.
 XX

CC The present sequence encodes one of 32 novel human secreted polypeptides.
 CC The nucleic acid molecules and polypeptides they encode may be used in
 CC the prevention, diagnosis and treatment of diseases such as
 CC immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus
 CC and human immuno-deficiency virus (HIV) infections), hyperproliferative
 CC disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases
 CC (e.g. scintular syndrome, Chaga's cardiomyopathy and coronary
 CC arteriosclerosis), angiogenic disorders (e.g. corneal graft
 CC neovascularisation and diabetic retinopathy), neurological disorders
 CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
 CC infectious diseases and/or for promoting wound healing, regeneration
 CC and/or chemotaxis. The nucleic acid molecules may be used to produce the
 CC secreted polypeptides. They may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acid
 CC sequences in samples. The polypeptides may be used as antigens in the
 CC production of antibodies and in assays to identify modulators of
 CC their expression and activity.
 CC
 XX
 SQ Sequence 1086 BP; 260 A; 316 C; 326 G; 184 T; 0 other;

Query Match 5.8%; Score 60; DB 22; Length 1086;
 Best Local Similarity 100.0%; Pred. No. 4.1e-12;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 978 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacctcgaggggggccc 1037
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1001 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacctcgaggggggccc 1060

RESULT 10
 AAF91929
 ID AAF91929 standard; cDNA; 1086 BP.
 XX
 AC AAF91929;
 XX
 DT 22-MAY-2001 (first entry)
 DE Human secreted protein-encoding gene 51 cDNA clone HAPQT56, SEQ ID NO:82.
 DE
 XX Human: secreted protein; proliferative disorder; cancer; tumour;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vulnery;
 KW cell culture; chemotaxis; food additive;
 KW binding partner identification; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200118022-A1.
 PD 15-MAR-2001.
 PD
 XX 31-AUG-2000; 2000WO-US24008.
 PF
 XX 03-SEP-1999; 99US-0152315.
 PR 03-SEP-1999; 99US-0152317.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX NI J, Baker KP, Birse CE, Fiscella M, Komatsoulis GA, Rosen CA;
 PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;
 PI Moore PA, Shi Y, Wei Y, Florence KA;
 XX
 XX WPI; 2001-203081/20.
 DR P-PDB; AAB87413.

PT Nucleic acid molecules encoding human secreted proteins, used in

Query Match 6.0%; Score 62; DB 19; Length 2307;
Best Local Similarity 100.0%; Pred. No. 7.1e-13;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgaaggggg 1035
DB 2216 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgaaggggg 2275

QY 1036 cc 1037
DB 2276 cc 2277

RESULT 8
AAV34270
ID AAV34270 standard; DNA: 737 BP.
XX
AC AAV34270;
XX
DT 28-JAN-1999 (first entry)
XX
DE Human secreted protein gene 63 clone HBJEM49.
XX
KW Human: secreted protein; fusion protein; gene therapy; protein therapy;
diagnosis; tissue: cancer; tumour; neurodegenerative disorder; leukaemia;
developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
PN MO9839446-A2.
XX
PD 11-SEP-1998.
XX
PF 06-MAR-1998; 98WO-US04492.
XX
XX 07-MAR-1997; 97US-0038621.
PR 07-MAR-1997; 97US-0040161.
PR 07-MAR-1997; 97US-0040162.
PR 07-MAR-1997; 97US-0040163.
PR 07-MAR-1997; 97US-0040333.
PR 07-MAR-1997; 97US-0040334.
PR 07-MAR-1997; 97US-0040336.
PR 11-APR-1997; 97US-0043311.
PR 11-APR-1997; 97US-0043312.
PR 11-APR-1997; 97US-0043313.
PR 11-APR-1997; 97US-0043314.
PR 11-APR-1997; 97US-0043315.
PR 11-APR-1997; 97US-0043568.
PR 11-APR-1997; 97US-0043569.
PR 11-APR-1997; 97US-0043576.
PR 11-APR-1997; 97US-0043578.
PR 11-APR-1997; 97US-0043580.
PR 11-APR-1997; 97US-0043669.
PR 11-APR-1997; 97US-0043670.
PR 11-APR-1997; 97US-0043671.
PR 11-APR-1997; 97US-0043672.
PR 11-APR-1997; 97US-0043674.
PR 23-MAY-1997; 97US-0047492.
PR 23-MAY-1997; 97US-0047500.
PR 23-MAY-1997; 97US-0047501.
PR 23-MAY-1997; 97US-0047502.
PR 23-MAY-1997; 97US-0047503.
PR 23-MAY-1997; 97US-0047581.
PR 23-MAY-1997; 97US-0047582.
PR 23-MAY-1997; 97US-0047583.
PR 23-MAY-1997; 97US-0047584.
PR 23-MAY-1997; 97US-0047585.

PR 23-MAY-1997; 97US-0047586.
PR 23-MAY-1997; 97US-0047587.
PR 23-MAY-1997; 97US-0047588.
PR 23-MAY-1997; 97US-0047589.
PR 23-MAY-1997; 97US-0047590.
PR 23-MAY-1997; 97US-0047592.
PR 23-MAY-1997; 97US-0047593.
PR 23-MAY-1997; 97US-0047594.
PR 23-MAY-1997; 97US-0047595.
PR 23-MAY-1997; 97US-0047596.
PR 23-MAY-1997; 97US-0047597.
PR 23-MAY-1997; 97US-0047598.
PR 23-MAY-1997; 97US-0047599.
PR 23-MAY-1997; 97US-0047600.
PR 23-MAY-1997; 97US-0047601.
PR 23-MAY-1997; 97US-0047612.
PR 23-MAY-1997; 97US-0047613.
PR 23-MAY-1997; 97US-0047614.
PR 23-MAY-1997; 97US-0047615.
PR 23-MAY-1997; 97US-0047617.
PR 23-MAY-1997; 97US-0047618.
PR 23-MAY-1997; 97US-0047632.
PR 23-MAY-1997; 97US-0047633.
PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048974.
PR 22-AUG-1997; 97US-0056630.
PR 22-AUG-1997; 97US-0056631.
PR 22-AUG-1997; 97US-0056632.
PR 22-AUG-1997; 97US-0056636.
PR 22-AUG-1997; 97US-0056637.
PR 22-AUG-1997; 97US-0056662.
PR 22-AUG-1997; 97US-0056664.
PR 22-AUG-1997; 97US-0056845.
PR 22-AUG-1997; 97US-0056862.
PR 22-AUG-1997; 97US-0056864.
PR 22-AUG-1997; 97US-0056872.
PR 22-AUG-1997; 97US-0056874.
PR 22-AUG-1997; 97US-0056881.
PR 22-AUG-1997; 97US-0056882.
PR 22-AUG-1997; 97US-0056884.
PR 22-AUG-1997; 97US-0056886.
PR 22-AUG-1997; 97US-0056887.
PR 22-AUG-1997; 97US-0056888.
PR 22-AUG-1997; 97US-0056889.
PR 22-AUG-1997; 97US-0056892.
PR 22-AUG-1997; 97US-0056893.
PR 22-AUG-1997; 97US-0056894.
PR 22-AUG-1997; 97US-0056903.
PR 22-AUG-1997; 97US-0056908.
PR 22-AUG-1997; 97US-0056909.
PR 22-AUG-1997; 97US-0056910.
PR 22-AUG-1997; 97US-0056911.
PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057761.
XX
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
PI Feng P, Fertile AM, Fischer CL, Graves KA, Greene JM, Hu JS;
PI Kyaw H, Latteur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
XX
XX WPI: 1998-609887/51.
XX P-PSDB: AAW75173.
XX
XX New isolated human genes and the secreted polypeptides they encode
XX - useful for diagnosis and treatment of e.g. cancers, neurological

CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SQ Sequence 1309 BP: 388 A: 321 C: 300 G: 300 T: 0 other;

Query Match 6.0%; Score 62; DB 22; Length 1309;
 Best Local Similarity 100.0%; Pred. No. 8.2e-13;
 Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgaaggagg 1035
 |||
 Db 1216 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgaaggagg 1275

QY 1036 cc 1037
 ||
 Db 1276 cc 1277

RESULT 6
 AAC59424
 ID AAC59424 standard; cDNA; 1547 BP.

AC AAC59424;
 XX
 DT 02-FEB-2001 (first entry)
 XX

DE Human secreted protein cDNA #33.
 XX

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anticancer;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; ss.
 XX

OS Homo sapiens.
 XX

PN W0200056765-A1.
 XX

PD 28-SEP-2000.
 XX

PF 16-MAR-2000; 2000WO-US06823.
 XX

PR 19-MAR-1999; 99US-0125364.
 XX

PR 08-DEC-1999; 99US-0169623.
 XX

PA (HUMA-) HUMAN GENOME SCI INC.
 XX

PI Rosen CA, Ruben SM, Komatsu N, et al.
 XX

DR WPI: 2000-602215/57.
 XX

P-PSDB: AAB33992.
 XX

PT Nucleic acid molecules encoding human secreted proteins, used in
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 PT Parkinson's diseases and cancers -
 XX

PS Claim 1: Page 347-348; 410pp; English.
 XX

CC The invention relates to the isolation of genes AAC59392-C59439 encoding
 CC 48 human secreted proteins AAB33963-B34006. The genes can be used to
 CC generate fusion proteins by linking to the gene for the human
 CC immunoglobulin G Fc portion (SEQID1) for increasing the stability of
 CC the fusion protein as compared to the human protein only. The genes and
 CC proteins are useful for preventing, ameliorating or treating medical
 CC conditions, e.g. by protein or gene therapy. The genes are isolated
 CC from a range of human tissues disclosed in the specification. The
 CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer, and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune

CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d)
 CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 XX

SQ Sequence 1547 BP: 410 A: 379 C: 337 G: 421 T: 0 other;

Query Match 6.0%; Score 62; DB 21; Length 1547;
 Best Local Similarity 100.0%; Pred. No. 7.8e-13;
 Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgaaggagg 1035
 |||
 Db 1483 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgaaggagg 1542

QY 1036 cc 1037
 ||
 Db 1543 cc 1544

RESULT 7
 AAV44318
 ID AAV44318 standard; cDNA to mRNA; 2307 BP.

AC AAV44318;
 XX

DT 07-OCT-1998 (first entry)
 XX

DE Maize starch-associated protein cDNA.
 XX

KW Starch-associated protein; maize; starch grain; soluble;
 KW transgenic plant; ss.
 XX

OS Zea mays.
 XX

PN Key Location/Qualifiers
 XX
 FT 33..1946
 FT CDS
 FT /tag=a
 FT /product="starch-associated protein"
 FT /note="no start codon given"
 XX

PN DE19653176-A1.
 XX

PD 25-JUN-1998.
 XX

PF 19-DEC-1996; 96DE-1053176.
 XX

PF 19-DEC-1996; 96DE-1053176.
 XX

PR 19-DEC-1996; 96DE-1053176.
 XX

PA (PLAN-) PLANTTEC BIOTECHNOLOGIE GMBH.
 XX

PI Emmermann M, Kossmann J;
 XX

DR WPI: 1998-349465/31.
 XX

P-PSDB: AAW64231.
 XX

PT DNA encoding maize starch-associated protein in potatoes - and
 PT corresponding protein, vectors, transgenic plants, antibody, starch,
 PT etc.
 XX

PS Claim 1: Page 43-46; 55pp; German.
 XX

CC This sequence encodes a novel maize starch-associated protein that
 CC occurs both bound to starch grains and in soluble form in plant cells.
 CC Such a protein can be used to produce transgenic plants or plant cells
 CC which are capable of synthesizing modified starch.
 XX

SQ Sequence 2307 BP: 693 A: 436 C: 555 G: 623 T: 0 other;

CC and cognitive disorders, schizophrenia, prostate diseases, obesity,
CC disorders involving osteoclasts such as osteoporosis, arthritis or
CC malignancies, diseases of testes, lung or thymus, digestive/endocrine
CC disorders, infections and AIDS. The polypeptides are also useful for
CC identifying their binding partners.
XX
XX Sequence 1018 BP: 251 A; 268 C; 280 G; 216 T; 3 other;

Query Match 6.0%; Score 62; DB 20; Length 1018;
Best Local Similarity 100.0%; Pred. No. 8,7e-13;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagggggg 1035
|||||
Db 924 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagggggg 983

OY 1036 cc 1037
||
Db 984 cc 985

RESULT 4

AAH34903
ID AAH34903 standard; CDNA: 1233 BP.

AC AAH34903:

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding CDNA SEQ ID NO:1985.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

KW colorectal carcinoma; ss.

OS Homo sapiens.

PN WO200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

DR WPI: 2001-235357/24.

DR P-PSDB; AAG75498.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,

PT useful for preventing, diagnosing and/or treating colorectal cancers -

XX Claim 1; Page 3486; 9803pp: English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the

CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX Sequence 1233 BP: 324 A; 286 C; 243 G; 369 T; 11 other;

Query Match 6.0%; Score 62; DB 22; Length 1233;
Best Local Similarity 100.0%; Pred. No. 8.3e-13;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagggggg 1035
|||||
Db 1142 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagggggg 1201

OY 1036 cc 1037
||
Db 1202 cc 1203

RESULT 5

AAH35025
ID AAH35025 standard; CDNA: 1309 BP.

AC AAH35025:

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding CDNA SEQ ID NO:2107.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

KW colorectal carcinoma; ss.

OS Homo sapiens.

PN WO200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

DR WPI: 2001-235357/24.

DR P-PSDB; AAG75620.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,

PT useful for preventing, diagnosing and/or treating colorectal cancers -

XX Claim 1; Page 3592; 9803pp: English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.

```

XX  MO9922243-A1.
PN
XX
PD  06-MAY-1999.
XX
XX  23-OCT-1998; 98WO-US22376.
XX
PR  24-OCT-1997; 97US-0063387.
PR  24-OCT-1997; 97US-0062784.
PR  24-OCT-1997; 97US-0063088.
PR  24-OCT-1997; 97US-0063089.
PR  24-OCT-1997; 97US-0063090.
PR  24-OCT-1997; 97US-0063091.
PR  24-OCT-1997; 97US-0063092.
PR  24-OCT-1997; 97US-0063097.
PR  24-OCT-1997; 97US-0063098.
PR  24-OCT-1997; 97US-0063099.
PR  24-OCT-1997; 97US-0063100.
PR  24-OCT-1997; 97US-0063101.
PR  24-OCT-1997; 97US-0063109.
PR  24-OCT-1997; 97US-0063110.
PR  24-OCT-1997; 97US-0063111.
PR  24-OCT-1997; 97US-0063148.
PR  24-OCT-1997; 97US-0063386.
XX
PA  (HUMA-) HUMAN GENOME SCI INC.
XX
PI  Brewer LA, Carter KC, Duan DR, Ebner R, Endress GA;
PI  Feng P, Florence C, Florence KA, Greene JM, Janat F;
PI  Kayw H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
PI  Ruben SM, Shi Y, Soppet DR, Wei Y, Young P;
XX
DR  WPI: 1999-303069/25.
DR  P-PSDB: AAY19590.
XX
PT  New isolated human genes and the secreted polypeptides they encode
XX
PS  Claim 3; Page 386-387; 546pp; English.
XX
CC  The specification describes cDNA sequences (AA61322-X61470) encoding
CC  human secreted proteins (AAY19442-Y19590). The polynucleotides and their
CC  corresponding secreted polypeptides are useful for preventing, treating
CC  or ameliorating medical conditions, e.g. by protein or gene therapy.
CC  Pathological conditions can also be diagnosed by determining the amount
CC  of the polypeptides in a sample or by determining the presence of
CC  mutations in the polynucleotides. Specific uses are described for each
CC  of the polynucleotides, based on which tissues they are most highly
CC  expressed in, and include developing products for the diagnosis or
CC  treatment of cancer, tumours, neurodegenerative disorders, developmental
CC  abnormalities and fetal deficiencies, blood disorders, leukemias,
CC  diseases of the immune system, autoimmune diseases, hepatic and renal
CC  disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's
CC  and cognitive disorders, schizophrenia, prostate diseases, obesity,
CC  disorders involving osteoclasts such as osteoporosis, arthritis or
CC  malignancies, diseases of testes, lung or thymus, digestive/endocrine
CC  disorders, infections and AIDS. The polypeptides are also useful for
CC  identifying their binding partners.
XX
SQ  Sequence 804 BP; 220 A; 199 C; 199 G; 183 T; 3 other:
XX
Query Match      6.0%; Score 62; DB 20; Length 804;
Best Local Similarity 100.0%; Pred. No. 9.2e-13;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY  976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactgcgagggggg 1035
    |||
DB  710 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactgcgagggggg 769
QY  1036 cc 1037
    ||
DB  770 cc 771

```

```

RESULT 3
AA61378
ID  AA61378 standard; cDNA; 1018 BP.
XX
AC  AA61378;
XX
DT  14-JUL-1999 (first entry)
XX
DE  DNA encoding a human secreted protein.
XX
KW  Human secreted protein; cancer; tumour; neurodegenerative disorder;
KW  developmental abnormality; fetal deficiency; blood disorder; leukemia;
KW  immune system disease; autoimmune disease; hepatic disease;
KW  renal disease; lymphoma; inflammation; allergy; ischemic shock;
KW  Alzheimer's; cognitive disorder; schizophrenia; prostate disease;
KW  obesity; osteoclast; osteoporosis; arthritis; malignancy; testis disease;
KW  lung disease; thymus disease; digestive disorder; endocrine disorder;
KW  infection; AIDS; ss.
XX
OS  Homo sapiens.
XX
XX  WO9922243-A1.
XX
PD  06-MAY-1999.
XX
XX  23-OCT-1998; 98WO-US22376.
XX
PR  24-OCT-1997; 97US-0063387.
PR  24-OCT-1997; 97US-0062784.
PR  24-OCT-1997; 97US-0063088.
PR  24-OCT-1997; 97US-0063089.
PR  24-OCT-1997; 97US-0063090.
PR  24-OCT-1997; 97US-0063091.
PR  24-OCT-1997; 97US-0063092.
PR  24-OCT-1997; 97US-0063097.
PR  24-OCT-1997; 97US-0063100.
PR  24-OCT-1997; 97US-0063109.
PR  24-OCT-1997; 97US-0063110.
PR  24-OCT-1997; 97US-0063111.
PR  24-OCT-1997; 97US-0063148.
PR  24-OCT-1997; 97US-0063386.
XX
PA  (HUMA-) HUMAN GENOME SCI INC.
XX
PI  Brewer LA, Carter KC, Duan DR, Ebner R, Endress GA;
PI  Feng P, Florence C, Florence KA, Greene JM, Janat F;
PI  Kayw H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
PI  Ruben SM, Shi Y, Soppet DR, Wei Y, Young P;
XX
DR  WPI: 1999-303069/25.
DR  P-PSDB: AAY19498.
XX
PT  New isolated human genes and the secreted polypeptides they encode
XX
PS  Claim 3; Page 341; 546pp; English.
XX
CC  The specification describes cDNA sequences (AA61322-X61470) encoding
CC  human secreted proteins (AAY19442-Y19590). The polynucleotides and their
CC  corresponding secreted polypeptides are useful for preventing, treating
CC  or ameliorating medical conditions, e.g. by protein or gene therapy.
CC  Pathological conditions can also be diagnosed by determining the amount
CC  of the polypeptides in a sample or by determining the presence of
CC  mutations in the polynucleotides. Specific uses are described for each
CC  of the polynucleotides, based on which tissues they are most highly
CC  expressed in, and include developing products for the diagnosis or
CC  treatment of cancer, tumours, neurodegenerative disorders, developmental
CC  abnormalities and fetal deficiencies, blood disorders, leukemias,
CC  diseases of the immune system, autoimmune diseases, hepatic and renal
CC  disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's

```


961	49	4.7	1178	21	AAC90469	Human uncoupling p
962	49	4.7	1180	20	AAV84347	Human Apo-2Dcr cDN
963	49	4.7	1181	19	AAV59803	Human secreted pro
964	49	4.7	1196	19	AAV20812	Human haematopoiet
965	49	4.7	1204	21	AAC59836	Human secreted pro
966	49	4.7	1210	21	AAZ50064	Imptiens balsamia
967	49	4.7	1211	21	AAV76729	Human ORX ORF284
968	49	4.7	1212	19	AAV59686	Human secreted pro
969	49	4.7	1215	20	AAV00620	Human secreted pro
970	49	4.7	1228	22	AAF33098	Human secreted pro
971	49	4.7	1233	20	AAF21869	Human breast and o
972	49	4.7	1240	22	AAZ21289	Human cDNA sequenc
973	49	4.7	1240	22	AAC91571	Human PRO3438 cDNA
974	49	4.7	1241	21	AAF21753	Human breast and o
975	49	4.7	1247	21	AAC59186	Human secreted pro
976	49	4.7	1249	21	AAZ04195	H. vulgare nicotia
977	49	4.7	1254	22	AAV05400	Human secreted pro
978	49	4.7	1255	20	AAV37455	Human secreted pro
979	49	4.7	1261	20	AAV04382	Human secreted pro
980	49	4.7	1274	21	AAF18072	Lung cancer associ
981	49	4.7	1276	17	AAV42071	Human nuclear prol
982	49	4.7	1285	16	AAV06744	TMMP-3 metalloprot
983	49	4.7	1293	21	AAV93129	Human secreted pro
984	49	4.7	1299	21	AAV16666	Human secreted pro
985	49	4.7	1302	12	AAQ11112	MHC class II antiq
986	49	4.7	1302	20	AAV90446	Human secreted pro
987	49	4.7	1307	19	AAV63176	cDNA from clone dx
988	49	4.7	1316	21	AAV95796	Apoptosis related
989	49	4.7	1316	22	AAV05658	Human secreted pro
990	49	4.7	1322	20	AAV04402	Human secreted pro
991	49	4.7	1322	21	AAV96798	Human pancreatic c
992	49	4.7	1335	21	AAV16099	Human prostate can
993	49	4.7	1337	20	AAV04345	Human secreted pro
994	49	4.7	1338	15	AAV05607	Rabbit zona pelluc
995	49	4.7	1338	20	AAZ22712	Rabbit zona pelluc
996	49	4.7	1338	20	AAV64789	Rabbit ZPC cDNA.
997	49	4.7	1338	21	AAZ95649	Rabbit zona pelluc
998	49	4.7	1338	21	AAZ46257	Rabbit oocyte zona
999	49	4.7	1338	21	AAZ33246	Rabbit zona pelluc
1000	49	4.7	1338	21	AAZ37801	Rabbit zona pelluc

ALIGNMENTS

RESULT 1

AAFS5866 standard; cDNA: 2087 BP.

AAFS5866:

17-APR-2001 (first entry)

Rat GLUTX1 coding sequence.

Rat: GLUTX: gene therapy; vaccine; hexose transport modulator; hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss; hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.

Rattus sp.

MO200104145-A2.

18-JAN-2001.

14-JUL-2000; 2000MO-1B01042.

14-JUL-1999; 99US-0143907.

27-AUG-1999; 99US-0151140.

23-FEB-2000; 2000US-0184285.

13-JUL-2000; 2000US-0616132.

(UYLA-) UNIV LAUSANNE.

XX Thorens B, Idberson M, Uldry M;
PI WPI: 2001-112615/12.
XX P-PSDB; AAB66933.
DR
XX
PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
XX e.g. ischemia and diabetes -
PS Claim 3; Page 71-73; 124pp; English.

XX The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and
XX AAB66932-AAB66941). The GLUTX proteins are related to the facultative
XX glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
XX function. The GLUTX proteins may be used in the diagnosis, prevention and
XX treatment of hexose transport disorders such as ischaemia, diabetes,
XX hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
XX neurodegenerative disease. The present sequence is the coding sequence
XX for rat GLUTX1.

SQ Sequence 2087 BP; 355 A; 673 C; 580 G; 479 T; 0 other;

Query Match 34.3%; Score 356; DB 22; Length 2087;
Best Local Similarity 100.0%; Pred. No. 1.6e-113;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 ctgggtgccatctccgcaagcctgctgattctacctgggctggcctggctgctgta 114
DB 1098 ctgggtgccatctccgcaagcctgctgattctacctgggctggcctggctgctgta 1157
QY 115 ggcagatgtgctctctcctcctgctgctgctgctgctgctgctgctgctgctgctgctc 174
DB 1158 ggcagatgtgctctcctcctcctgctgctgctgctgctgctgctgctgctgctgctgctc 1217
QY 175 ctcatgtccagagatcttcctcctcctcctcctcctcctcctcctcctcctcctcctc 234
DB 1218 ctcatgtccagagatcttcctcctcctcctcctcctcctcctcctcctcctcctcctc 1277
QY 235 acccaatgttcatatgctcttctggtgacaaagaattaaacagcatatgagatccctc 294
DB 1278 acccaatgttcatatgctcttctggtgacaaagaattaaacagcatatgagatccctc 1337
QY 295 agaccctacggcgcccttcctgctcaccgctgctccttcctgctatccctcctcctcctc 354
DB 1338 agaccctacggcgcccttcctgctcaccgctgctccttcctgctatccctcctcctcctc 1397
QY 355 ctcaactgttcctcctgagactaaagcagagactctgaaacaatcacagccattt 410
DB 1398 ctcaactgttcctcctgagactaaagcagagactctgaaacaatcacagccattt 1453

RESULT 2

AAK61470 standard; cDNA: 804 BP.

AAK61470:

14-JUL-1999 (first entry)

DNA encoding a human secreted protein.

Human secreted protein; cancer; tumor; neurodegenerative disorder; developmental abnormality; fetal deficiency; blood disorder; leukemia; immune system disease; autoimmune disease; hepatic disease; renal disease; lymphoma; inflammation; allergy; ischemic shock; Alzheimer's; cognitive disorder; schizophrenia; prostate disease; obesity; osteoclast; osteoporosis; arthritis; malignancy; testes disease; lung disease; thymus disease; digestive disorder; endocrine disorder; infection; AIDS; ss.

Homo sapiens.

C 231	54	5.2	1500	20	AAAX77395	Murine C12.2bs pol
C 232	54	5.2	1509	21	AAA53599	M. tuberculosis an
233	54	5.2	1517	21	AAA26334	Human secreted pro
234	54	5.2	1554	20	AAAX16675	Xenopus W545 prot
235	54	5.2	1556	18	AAAT94535	Maize cinnamoyl Co
236	54	5.2	1568	18	AAAT94534	Alfatia cinnamoyl
237	54	5.2	1585	19	AAV47581	Leishmania antigen
238	54	5.2	1586	14	AAQA4967	Rape acyl-ACP thio
239	54	5.2	1605	19	AAV59687	Human secreted pro
240	54	5.2	1607	22	AAAD05431	Human secreted pro
241	54	5.2	1637	21	AAAF16335	Human prostate can
242	54	5.2	1638	17	AAAT06480	Cystathionine gamm
243	54	5.2	1643	22	AAAD07753	Human secreted pro
244	54	5.2	1649	22	AAAD07600	Human secreted pro
245	54	5.2	1663	21	AAAC59302	Human secreted pro
246	54	5.2	1693	22	AAAD07724	Human secreted pro
247	54	5.2	1707	22	AAAF72809	Secreted protein g
248	54	5.2	1710	14	AAQA7966	Rape acyl-ACP thio
249	54	5.2	1712	21	AAAF16208	Human prostate can
250	54	5.2	1735	21	AAAF8385	Human secreted pro
251	54	5.2	1737	22	AAAF29901	Human h2252 DNA.
252	54	5.2	1901	21	AAAF7405	1-deoxy-D-xylose
253	54	5.2	1968	21	AAAC69646	Human secreted pro
254	54	5.2	1992	22	AAAD05341	Human secreted pro
255	54	5.2	1996	21	AAAC77813	Human cancer assoc
256	54	5.2	2012	21	AAAC59942	Human secreted pro
257	54	5.2	2034	22	AAAH33087	Human colon cancer
258	54	5.2	2040	18	AAAT62137	Human secreted pro
259	54	5.2	2040	19	AAV47557	Leishmania brazili
260	54	5.2	2040	19	AAV47557	Leishmania antige
261	54	5.2	2065	21	AAV50436	Human Casper gene.
262	54	5.2	2163	22	AAAD00808	Human irritable bo
263	54	5.2	2229	22	AAAF91900	Human secreted pro
264	54	5.2	2233	19	AAAH34046	Human colon cancer
265	54	5.2	2248	21	AAAF16340	Human retinoid rec
266	54	5.2	2248	22	AAAD08448	Human prostate can
267	54	5.2	2248	22	AAAH32575	Human secreted pro
268	54	5.2	2278	21	AAAB89962	Corn ADA2 transcr
269	54	5.2	2429	21	AAAH35009	Human colon cancer
270	54	5.2	2443	22	AAAF45088	Human secreted pro
271	54	5.2	2462	21	AAAC69525	Human secreted pro
272	54	5.2	2503	15	AAAO53480	PNP30 xylanase CD
273	54	5.2	2503	16	AAAT05571	MEKK2 CDNA. Mus s
274	54	5.2	2503	19	AAV22677	CDNA encoding a ml
275	54	5.2	2503	20	AAV22681	CDNA encoding a ml
276	54	5.2	2503	20	AAZ31878	Mitogen ERK kinase
277	54	5.2	2503	20	AAZ31882	MEKK2 protein codi
278	54	5.2	2503	20	AAV45615	Murine MEKK2 codi
279	54	5.2	2503	21	AAAH92227	Murine MEKK6 codi
280	54	5.2	2541	20	AAAX04343	Human secreted pro
281	54	5.2	2583	21	AAAZ38752	Soybean 1-deoxy-D-
282	54	5.2	2668	20	AAZ07196	Human lung tumour-
283	54	5.2	2668	21	AAAC79150	Human lung tumour-
284	54	5.2	2751	22	AAAF91911	Human secreted pro
285	54	5.2	2775	18	AAAT72271	Human B cell surfa
286	54	5.2	2775	21	AAAO8399	Human RPI05 nucleo
287	54	5.2	2853	22	AAAF33096	Human secreted pro
288	54	5.2	2865	22	AAAF33259	Human secreted pro
289	54	5.2	2908	21	AAAC3384	Human secreted pro
290	54	5.2	2908	22	AAAS03899	Human secreted pro
291	54	5.2	3066	19	AAV09025	Homo sapiens 20q13
292	54	5.2	3089	16	AAAT05572	MEKK3 CDNA. Mus s
293	54	5.2	3089	19	AAZ22678	CDNA encoding a mi
294	54	5.2	3089	20	AAV231879	Mitogen ERK kinase
295	54	5.2	3089	20	AAV45616	MEKK3 protein codi
296	54	5.2	3109	21	AAAH92224	Murine MEKK3 codi
297	54	5.2	3109	22	AAAH35019	Human colon cancer
298	54	5.2	3152	21	AAAC59088	Human secreted pro
299	54	5.2	3224	22	AAAH19185	Human secreted pro
300	54	5.2	3224	22	AAAH19185	Human lung tumour
301	54	5.2	3375	20	AAZ07192	Human lung tumour
302	54	5.2	3375	20	AAZ07192	Human lung tumour
303	54	5.2	3375	20	AAZ07192	Human lung tumour

QY 1 tggcggccgctctagaactagtgatccccgggctgcaggaattcgacagact 56
|||||
DB 7 TGGCGGCCCTCTAGAACTAGTGATCCCCGGGCTGCAGGAATTCGCACGAGCT 62

RESULT 14

165496 1601 bp DNA PAT 07-OCT-1997
LOCUS Sequence 6 from patent US 5667997.
DEFINITION 165496
ACCESSION 165496.1 GI:2482066
VERSION 165496.1 GI:2482066
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1601)
AUTHORS Voelker,T,Alois,Davies,H,Maelor and Knutzon,D.
TITLE C8 and C10 medium-chain thioesterases in plants
JOURNAL Patent: US 5667997-A 6 16-SEP-1997;
FEATURES Location/Qualifiers
Source 1..1601
/organism="unknown"

BASE COUNT 375 a 372 c 429 g 425 t
ORIGIN

Query Match 5.4%; Score 56; DB 6; Length 1601;
Best Local Similarity 100.0%; Pred. No. 7.7e-21;

Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggcggccgctctagaactagtgatccccgggctgcaggaattcgacagact 56
|||||
DB 7 TGGCGGCCCTCTAGAACTAGTGATCCCCGGGCTGCAGGAATTCGCACGAGCT 62

RESULT 15

A82595 1711 bp DNA Circular PAT 21-JAN-2000
LOCUS A82595
DEFINITION Sequence 3 from Patent WO9854316.
ACCESSION A82595
VERSION A82595.1 GI:6732340
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1711)
AUTHORS Graham,D., and Renard,S.
TITLE SODIUM CHANNEL RECEPTOR
JOURNAL Patent: WO 9854316-A 3 03-DEC-1998;
GRAHAM DAVID (FR); RENARD STEPHANE (FR)

FEATURES

Source 1..1711
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="CEREBELLUM"

BASE COUNT 329 a 578 c 481 g 323 t
ORIGIN

Query Match 5.4%; Score 56; DB 6; Length 1711;
Best Local Similarity 100.0%; Pred. No. 7.7e-21;

Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggcggccgctctagaactagtgatccccgggctgcaggaattcgacagact 56
|||||
DB 18 TGGCGGCCCTCTAGAACTAGTGATCCCCGGGCTGCAGGAATTCGCACGAGCT 73

AUTHORS Imanishi, S.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-1997) to the DDBJ/EMBL/GenBank databases.
Shunshuke Imanishi, School of Agriculture, Nagoya University,
Laboratory of Biochemistry, Furo-cho, Chikusa, Nagoya 464-01, Japan
(E-mail: i45184@nuc.cc.nagoya-u.ac.jp, Tel: 052-789-4098,
Fax: 052-789-4094)
2 (sites)
REFERENCE
AUTHORS Kojima, H., Hashizume, K., Imanishi, S. and Nakamura, K.
TITLE Early Jasmonate-Inducible genes of Tobacco Cells
JOURNAL Unpublished (1997)
FEATURES
Location/Qualifiers
1..518
/organism="Nicotiana tabacum"
/strain="bright yellow 2"
/db_xref="taxon:4097"
<1..>518
/codon_start=2
/product="HYJ15"
/protein_id="BA021615.1"
/db_xref="GI:2280518"
/translation="LELRGGGRSRTSGPGIOEFGTSSTFENRPTGHSNGRIIDF
ITQSFKLPFLNAYLDGAAFTQGVNFVAGGALTALTSMTARNLPLTPNPLAOLG
WFKTHLOSTCGSCRCADSLKNSLIWMGEWGNIDYNGFPQNMQISEVRYVNVVAGIM
RCIKRCDARSRH"

BASE COUNT 121 a 142 c 122 g 133 t
ORIGIN

Query Match 5.4%; Score 56; DB 8; Length 518;
Best Local Similarity 100.0%; Pred. No. 8.3e-21;
Matches 56: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tggcgccgcctctagaactagtagtgcacccggcgctgcaggaattcgcagcagact 56
|||||
Db 19 TGGCGCCGCTCTAGACTAGTAGTGCACCCGCGCTGCAGCAATTGCGCAGCAGCT 74

RESULT 11
AB002449 1228 bp mRNA PRI 13-FEB-1999
LOCUS Homo sapiens mRNA from chromosome 5q21-22, clone:843EX.
DEFINITION
AB002449
VERSION AB002449.1 GI:2943812
KEYWORDS advanced lung cancer.
SOURCE Homo sapiens tissue_l1b:fetal brain cDNA to mRNA, clone:843EX.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1228)
REFERENCE
AUTHORS Ueno, K.
JOURNAL Direct Submission
Submitted (29-MAR-1997) to the DDBJ/EMBL/GenBank databases.
Kiyonobu Ueno, Osaka University Medical School, Department of
Medicine III; 2-2, Yamada-oka, Suita, Osaka, Japan, Suita, Osaka
565, Japan (E-mail: ueno-kei@ichu.hosp.go.jp, Tel: 06-879-3835,
Fax: 06-879-3839)
2 (sites)
REFERENCE
AUTHORS Ueno, K., Kunagai, T., Kijima, T., Kishimoto, T. and Hosoe, S.
TITLE Cloning and tissue expression of cDNAs from chromosome 5q21-22
JOURNAL Hum. Genet. 102 (1), 63-68 (1998)
MEDLINE
98141129
FEATURES
Location/Qualifiers
1..1228
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="843EX"
/map="5q21-22"
/tissue_l1b="fetal brain"
<1..>1228
/note="expressed ubiquitously"

BASE COUNT 415 a 225 c 205 g 377 t 6 others
ORIGIN

Query Match 5.4%; Score 56; DB 9; Length 1228;
Best Local Similarity 100.0%; Pred. No. 7.9e-21;
Matches 56: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tggcgccgcctctagaactagtagtgcacccggcgctgcaggaattcgcagcagact 56
|||||
Db 31 TGGCGCCGCTCTAGAACTAGTAGATCCCGGCGCTGCAGCAATTGCGCAGAGCT 86

RESULT 12
AX101122 1388 bp DNA PAT 10-APR-2001
LOCUS Sequence 1 from Patent WO0121650.
DEFINITION
AX101122
ACCESSION AX101122
VERSION AX101122.1 GI:13619973
KEYWORDS
SOURCE Dimorphotheca.
ORGANISM Dimorphotheca.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; eunasterids II; Asterales; Asteraceae; Asteroideae;
Calenduleae.
1 (bases 1 to 1388)
REFERENCE
AUTHORS Coldren, C., Flint, D., Hallahan, D.L. and Wang, H.
TITLE Cis-prenyltransferases from plants
JOURNAL Patent: WO 0121650-A 1 29-MAR-2001;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
Location/Qualifiers
1..1388
/organism="Dimorphotheca"
/db_xref="taxon:13490"

BASE COUNT 437 a 268 c 297 g 386 t
ORIGIN

Query Match 5.4%; Score 56; DB 6; Length 1388;
Best Local Similarity 100.0%; Pred. No. 7.8e-21;
Matches 56: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tggcgccgcctctagaactagtagtgcacccggcgctgcaggaattcgcagcagact 56
|||||
Db 231 TGGCGCCGCTCTAGACTAGTAGATCCCGGCGCTGCAGCAATTGCGCAGCAGCT 286

RESULT 13
I14842 1601 bp DNA PAT 02-APR-1996
LOCUS Sequence 10 from patent US 5455167.
DEFINITION
I14842
ACCESSION I14842
VERSION I14842.1 GI:1249751
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
1 (bases 1 to 1601)
REFERENCE
AUTHORS Voelker, T.A. and Davies, H.M.
TITLE Medium-chain thioesterases in plants
JOURNAL Patent: US 5455167-A 10 03-OCT-1995;
MEDLINE
98141129
FEATURES
Location/Qualifiers
1..1601
/organism="unknown"

BASE COUNT 375 a 372 c 429 g 425 t
ORIGIN

Query Match 5.4%; Score 56; DB 5; Length 1601;
Best Local Similarity 100.0%; Pred. No. 7.7e-21;
Matches 56: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/product="myosin heavy chain"
/protein_id="AAB29713.2"
/db_xref="GI:8250661"
/translation="PPGCRNSARGCLLEGLDLAEOETWVENDKQOOLDEKLKKPEFM
/translating="PPGCRNSARGCLLEGLDLAEOETWVENDKQOOLDEKLKKPEFM
SRLKSLKEDLQALGMOLOKRLKELQARLEELPELEASRAEAKAKRQSLRLEAE
ISERLEEGAGTSAQIENKKREAFQKMRDLLEATIQHATATLTKKADSAEL
GEQIDNLRVOKLEKSEKMETIDDLASMEVTSKSKNLEKCRILEOVSELK
KEEDQRLINELTAQRGLTQESGEYSQDLDEKSLVSLSGKAQFTQOILEKROL
EEVYAKSALAHALQSSRHQDLLEQYEEVAVAKASLEKTKORLONEVEDMIVERT
DAIOTRELEBAPKKAQRLQDAEHEVAVAKASLEKTKORLONEVEDMIVERT
NACAAADKKORNFDKILAEKQYEEVAVAKASLEKTKORLONEVEDMIVERT
OLPILKRNKLOOEISDTPEOIAEGGRHLEKIKKQIROPKSELQALAEPAEASL
EHEGKILRTLELNOYKSELDRTKIAEDDEIDOKRNHIVESMOSTLDAEISRN
DAIRIKKMEGDLNEMETOLNHSNNMAEALRNFTQGLIKDTQLHLDALRQEDL
KEOLAMERRANLQAELELRLATLEQTERSKIAEOLIDASERVOLHTRQNTSLN
TKKLETDISQIOGEMEDIVQEARNAEKAKATIDAMAEELKEDQTSANLEPMK
KNLQTYVDLQRLDEAQLKGGKQIOLLEARVRELEGEVEEOKRNYEALIKGLR
KHEERVKELTYQTEDEKRNVLRIQDLYDKLSKVKAKYKROAEAEQSNVLAERFKI
QHELEAEERADIAESOVNKLKRSREYHTKITSBE"

BASE COUNT 1005 a 650 c 871 g 398 t
ORIGIN

Query Match 6.0%; Score 62; DB 10; Length 2924;
Best Local Similarity 100.0%; Pred. No. 2,7e-24;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 976 aaactcgagggg 1035
|||||
Db 2859 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagggg 2918

OY 1036 cc 1037
||
Db 2919 cc 2920

RESULT 7
LOCUS AR070327 2605 bp DNA PAT 18-FEB-2000
DEFINITION Sequence 4 from patent US 5892010.
ACCESSION AR070327
VERSION AR070327.1 GI:7221215
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2605)
AUTHORS Gray,J., Collins,C., Hwang,S., Godfrey,T., Kowbel,D. and Rommens,J.
TITLE Genes from the 20013 amplicon and their uses
JOURNAL Patent: US 5892010-A 4 06-APR-1999;
FEATURES
source 1..2605
location/Qualifiers
BASE COUNT 698 a 625 c 550 g 732 t
ORIGIN

Query Match 5.6%; Score 58; DB 6; Length 2605;
Best Local Similarity 100.0%; Pred. No. 5.4e-22;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tggcgccgctctagaactagatcccccggcgctgcaggaattcgacagactg 58
|||||
Db 55 tggcgccgctctagaactagatcccccggcgctgcaggaattcgacagactg 112

RESULT 8
LOCUS AR060385 2085 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 7 from patent US 5840568.
ACCESSION AR060385
VERSION AR060385.1 GI:5986835
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2085)
AUTHORS Pfrendenschuh,M.
TITLE Hodgkin's disease associated molecules and uses thereof
JOURNAL Patent: US 5840568-A 7 24-NOV-1998;
FEATURES
source 1..2085
location/Qualifiers
BASE COUNT 746 a 398 c 557 g 384 t
ORIGIN

Query Match 5.5%; Score 57; DB 6; Length 2085;
Best Local Similarity 100.0%; Pred. No. 2e-21;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tggcgccgctctagaactagatcccccggcgctgcaggaattcgacagactg 57
|||||
Db 16 tggcgccgctctagaactagatcccccggcgctgcaggaattcgacagactg 72

RESULT 9
LOCUS AX062918/c 380 bp DNA PAT 24-JAN-2001
DEFINITION Sequence 545 from Patent WO0100828.
ACCESSION AX062918
VERSION AX062918.1 GI:12540776
KEYWORDS
SOURCE human.
ORGANISM human.
REFERENCE 1 (bases 1 to 380)
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Velivick,T.S.,
Carter,D., Reiter,M.W. and Mannion,J.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0100828-A 545 04-JAN-2001;
CORIXA CORPORATION (US)
FEATURES
source 1..380
location/Qualifiers
misc_feature 1..380
/db_xref="taxon:9606"
/organism="Homo sapiens"
BASE COUNT 100 a 76 c 102 g 97 t 5 others
ORIGIN

Query Match 5.4%; Score 56; DB 6; Length 380;
Best Local Similarity 100.0%; Pred. No. 8.4e-21;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tggcgccgctctagaactagatcccccggcgctgcaggaattcgacagactg 56
|||||
Db 332 tggcgccgctctagaactagatcccccggcgctgcaggaattcgacagactg 277

RESULT 10
LOCUS AB005878 518 bp mRNA PLN 25-JUL-1997
DEFINITION Nicotiana tabacum mRNA for BY15, partial cds.
ACCESSION AB005878
VERSION AB005878.1 GI:2280517
KEYWORDS BY15.
SOURCE Nicotiana tabacum (strain:bright yellow 2) cDNA to mRNA.
ORGANISM Nicotiana tabacum
REFERENCE 1 (bases 1 to 518)
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.

SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE	Novel Human cDNA clones with function of inhibiting cancer cell growth
JOURNAL	unpublished
REFERENCE	2 (bases 1 to 1977)
AUTHORS	Gu,J.R., Wan,D.F., Zhao,X.T., Zhou,X.M., Jiang,H.Q., Zhang,P.P., Qin,W.X., Huang,Y., Qiu,X.K., Qian,L.F., He,L.P., Li,H.N., Yu,Y., Yu,J., and Han,L.H.
TITLE	Direct Submission
JOURNAL	Submitted (21-DEC-1999) National Laboratory for Oncogenes & Related Genes, Shanghai Cancer Institute, 25 Ln 2200, Xie-Tu Road, Shanghai 200032, People's Republic of China
FEATURES	source
CDS	1..1977 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="pp1746" 817..1878 /codon_start=1 /product="unknown" /protein_id="AAGI7236.1" /db_xref="GI:10441918" translation="MHNVFFKKHYMTFECCNLTGCEFPATIGKTRQWVAPKKKKKKK KKKKKSRGGPPYGNLLPALISGGPFRHTDGLGASARGISPRLHPGVARGGGGR ORAHVAARGCGHPGILWLRLPGAAGMADLGGLGGQVGAGCGDALDPASAAT TCHPGOCNNVTLGNAHMGYEFTVACAGACAPSMHRSQSVSHMTNTRTIDPEILESS RYPYLRRFELRRSGRGRRFGSDGYTLRELPLEEALLSVLTRAFRPYGLHGGERP GARLLNLKRNGRTVMLGKSTSYTVPGDYFCILTTGGCGYGDEPDAPPSPPSPOA LAFFEHSVEYERAAOKAV"
BASE COUNT	420 a 608 c 605 g 344 t
ORIGIN	
Query Match	6.1%; Score 63; DB 9; length 1977;
Best Local Similarity	100.0%; Pred. NO. 7.5e-25;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 975	caaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacacgcgaqgqgg 1034 Db 920 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCTCGAGCGGGG 979
QY 1035 gcc 1037 	
Db 980 GCC 982	
RESULT 5	
LOCUS A91160 2307 bp DNA PAT 22-JAN-2000	
DEFINITION Sequence 5 from Patent WO9827212.	
ACCESSION A91160	
VERSION A91160.1 GI:6740195	
KEYWORDS Zea mays. Zea mays. Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.	
SOURCE ORGANISM	Zea mays. Zea mays. Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE	1 (bases 1 to 2307) Emmertmann,M. and Kossmann,J. NOVEL NUCLEIC ACID MOLECULES FROM MAIZE AND THEIR USE FOR THE PRODUCTION OF MODIFIED STARCH Patent: WO 9827212-A 5 25-JUN-1998:
AUTHORS	EMMERMANN MICHAEL (DE); KOSSMANN JENS (DE)
JOURNAL	location/Qualifiers
FEATURES	I..2307

CDS	/organism="Zea mays" /db_xref="taxon:4577" <33.1946 /note="unnamed protein product" /codon_start=1 /protein_id="CAB69412.1" /db_xref="GI:6740196" /translation="HELEGLEAVELRPLLDSREMKDLIFLDIALDSTFPAIER SYELNDAAPEKIMVFLSVLENLALSIDNEDILVLCIKGNQALEMAKOKDDQMALV AKAFLDNRKRLASKGEQYHNMMQPSAYLSLISIDQAVNIPTKEIIRGSAATLS ALAIRPPEVLNHLNLSMOWI SPEVSGYVAVDELAVONKSYDKPTILVAKSVGR EELIDPDGVVGIPTDMPDVLSHVSRANRSKVLFAFCDHTTSLSELYGDKLFSPR TSALITRETFSELOOSSPNAEVAHVPISLAKKFLKVAISABESEBVGAK SRNIATLKGEPSWVGVPVTAIPFGFEKVLGSLGNKEVAQSTIEKLTIRLAQDFSA LGEIRKAVLNLITAPMQLVNELEKRMILGSGMPDDEDDKRMEOQAMAIKVMASRMB RAYSTRKVLKIDHEYLISMAVLQVEVNAADYAFVHTNPSSGDSSEIYAEVYKLGCTE LVAGYSGRAMSFVCKKDDLDSPKLLGSPSKIGLFTROSIIFRSDSGEDLEGVAGAG LYSDVPEDEVEVLYDVTPTLVDKPRFSSILSSIRAGHAIEELVXGSDVEGVYK DGKIVYQTRPROM"	
BASE COUNT	693 a	436 c 555 g 623 t
ORIGIN		
Query Match	6.0%	Score 62: DB 6: Length 2307:
Best Local Similarity	100.0%	Prod. No. 2.8e-24:
Matches	62: Conservative	0: Mismatches 0: Indels 0: Gaps 0:
OY	976	aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacctcgaaggagg 1035
Db	2216	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTCAGAGGGGG 2275
OY	1036	cc 1037
Db	2276	cc 2277
RESULT	6	
LOCUS	S68736	2924 bp mRNA ROD 05-JUN-2000
DEFINITION	Rattus sp. myosin heavy chain mRNA, partial cds.	
ACCESSION	S68736	
VERSION	S68736.1	GI:544938
KEYWORDS		
SOURCE	Rattus sp.	
ORGANISM	Rattus sp.	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. 1 (bases 1 to 2924) Ogata,I., Saez,C.G., Greenwel,P., Ponce M de,L., Geerts,A., Leinwand,L.A. and Rojkind,M. Rat liver fat-storing cell lines express sarcomeric myosin heavy chain mRNA and protein Cell Motil. Cytoskeleton 26 (2), 125-132 (1993) 94116097 GenBank staff at the National Library of Medicine created this entry [NCBI gidsq 1436571] from the original journal article. This sequence comes from Fig. 3. Location/Qualifiers 1..2924 /organism="Rattus sp." /db_xref="taxon:10118" /cell_line="CC14" /cell_type="fat storing cells" /tissue_type="clithotic liver" 1..2924 /gene="myosin heavy chain, MHC" <1..2679 /gene="myosin heavy chain, MHC" /note="MHC: This sequence comes from Fig. 3: conceptual translation presented here differs from translation in publication" /codon_start=1	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 20:05:17 ; Search time 4552.9 Seconds
(without alignments)
3757.512 Million cell updates/sec

Title: US-09-516-493-9

Perfect score: 1037
Sequence: 1 tggcgcgcctctagaacta.....aaaactcgaaggcgc 1037

Scoring table:

OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pt:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_inv:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	356	34.3	2087	6	AX076669
2	356	34.3	2087	10	RN0245935
3	356	34.3	2189	10	AB033418
4	63	6.1	1977	6	AF217994
5	62	6.0	2307	6	A91160
6	62	6.0	2307	6	A91160
7	58	5.6	2605	6	AR070327
8	57	5.5	2085	6	AR060385
9	56	5.4	380	6	AX062918
10	56	5.4	518	8	AB005878
11	56	5.4	1228	6	AB002449
12	56	5.4	1388	6	AX101122
13	56	5.4	1601	6	114842
14	56	5.4	1601	6	165496
15	56	5.4	1711	6	AB2595
16	56	5.4	1711	6	AB2786
17	56	5.4	1711	6	AB2786
18	55	5.3	2876	8	CS1011095
19	55	5.3	912	3	TE286120
20	55	5.3	1190	6	AR087649
21	55	5.3	1462	10	BC006018
22	55	5.3	1585	8	OSCHINDPR
23	55	5.3	1605	5	AX027925
24	55	5.3	1701	9	AR028859
25	55	5.3	1798	8	ATCYC2B
26	55	5.3	2418	6	AX063691
27	55	5.3	2448	6	A93016
28	55	5.3	15528	12	PEAVGEN
29	55	5.2	15528	6	AR068753
30	54	5.2	289	6	163077
31	54	5.2	289	6	163077
32	54	5.2	515	10	MMPCG60
33	54	5.2	531	8	AB005879
34	54	5.2	536	6	AR068751
35	54	5.2	536	6	163075
36	54	5.2	562	3	DMELAST
37	54	5.2	824	6	PAB131096
38	54	5.2	1387	6	AX047063
39	54	5.2	1410	9	BC008364
40	54	5.2	1486	8	AF157496
41	54	5.2	1504	8	AT1134T
42	54	5.2	1544	5	OLANEXM3
43	54	5.2	1554	5	AP065135
44	54	5.2	1556	6	A62300
45	54	5.2	1556	6	AR145428
46	54	5.2	1568	6	A62298
47	54	5.2	1568	6	AR145427
48	54	5.2	1585	6	AX134055
49	54	5.2	1586	6	A76337
50	54	5.2	1586	6	BNPML3
51	54	5.2	1639	6	AR072149
52	54	5.2	1710	6	A76335
53	54	5.2	1710	6	BNPML2
54	54	5.2	1725	8	AF007785
55	54	5.2	1731	10	AY012159
56	54	5.2	2040	6	AR054175
57	54	5.2	2040	6	AX134006
58	54	5.2	2045	6	AR156902
59	54	5.2	2046	8	AF060519
60	54	5.2	2062	8	WHTEFLX
61	54	5.2	2101	9	BC000090
62	54	5.2	2186	3	PVA250829
63	54	5.2	2233	6	192592
64	54	5.2	2303	6	AR008277
65	54	5.2	2303	6	AR008281
66	54	5.2	2303	6	AR068731
67	54	5.2	2303	6	AR085073
68	54	5.2	2303	6	AR085077
69	54	5.2	2503	6	AR098266
70	54	5.2	2503	6	AR098270

AX076669 Sequence
A745935 Rattus no
AB033418 Rattus no
AF217994 Homo sapi
A91160 Sequence 5
S68736 Rattus sp.
AR070327 Sequence
AR060385 Sequence
AX062918 Sequence
AB005878 Nicotiana
AB002449 Homo sapi
AX101122 Sequence
114842 Sequence 10
165496 Sequence 6
AB2595 Sequence 3
AB2786 Sequence 1
A1011095 Citrus si
Z86120 T. evansi MR
AR087649 Sequence
AB4916 Sequence 7
BC006018 Mus muscu
Z54153 O. sativa MR
AX027925 Sequence
AB028859 Homo sapi
Z31401 A. thaliana
AX063691 Sequence
A132110 Oryctolag
A93016 Sequence 4
Y07862 Cloning vec
AR068753 Sequence
163077 Sequence 3
Y11505 M. musculus
AB005879 Nicotiana
AR068751 Sequence
163075 Sequence 1
X85998 D. melanogas
A1131096 Picea abi
AX047063 Sequence
BC008364 Homo sapi
AF157496 Arabidops
A1011753 Arabidops
Y11254 O. latipes m
AR065135 Xenopus 1
A62300 Sequence 3
AR145428 Sequence
A62298 Sequence 1
AR145427 Sequence
AX134055 Sequence
A76337 Sequence 3
X73850 B. napus (PN
AR072149 Sequence
A76335 Sequence 1
X73849 B. napus (PN
AF007785 zea mays
AY012159 Mus muscu
AR054175 Sequence
AX134006 Sequence
AR156902 Sequence
AF060519 Cuphea ho
M90077 Wheat trans
BC000090 Homo sapi
A7250829 Penaeus v
192592 Sequence 1
AR008277 Sequence
AR008281 Sequence
AR068731 Sequence
AR085073 Sequence
AR085077 Sequence
AR098266 Sequence
AR098270 Sequence

Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project. No 51 sequence available. This clone (DKR2p434P0816) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers

FEATURES

Source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IDRF2p434P0816"
/clone.lib="434 (synonym: htes3)"
/tissue.type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/notes="vector: pSport1; Site_1: NotI; Site_2: SalI"
125 c 103 g 82 t

```

BASE COUNT	79 a	125 c	103 g	82 t
ORIGIN				

ORIGIN

Query Match	Score	DB	Length
21.08	380	10	389
100.00	380	10	389

Best Local Similarity 100.0%; Pred. No. 1.2e-87;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1278	tcctgacattcgaagtcctcttcaactgtctctgtgtctccctgaactaaagaagaagctc	1337
Db	1	tcctgacattcctcagtcgctcctttttacattttgtctgtctccctcaaaactaaagaagaagctc	60
QY	1338	tggaacaatatacaagcccatcttttgagggcgatgatacaagcactcaactagggatgagca	1397
Db	61	tggaacaataatcacagcccatctttttgaggggggatgatacagccacttcactgtaggagca	120
QY	1398	agctctgtgactccaagcttgggcccgaagcccagaagccctgtcctgcgcccaaggagagca	1457
Db	121	agcctgtgactccaagcttggggcccagaagccagagccctgcctgcgcccaaggagaccca	180
QY	1458	atccagagcccttgagagccttggctctgagaggtccctccctctctgtgcatgctccctcagc	1517
Db	181	atccagagccctcttgagagccttggctctgagaggtccctccctctctgtgcatgctccctcagc	240
QY	1518	ccatgaccccgaggcctagagaggtcactgtcctcgtcttccagctcctgtcgtgctcctag	1577
Db	241	ccatgacaccccgaggccttagagaggtcactgtcctcgtcttccagctcctgtcgtgctcctag	300
QY	1578	gactccaggaacaccttcagacttttcagaagccttgcagagccttcacatgagcaagactaa	1637
Db	301	gactccaggaacacaccttcagacttttcagaagccttgcagagccttcacatgagcaagactaa	360
QY	1638	agcagcggaagagagtgatg	1657
Db	361	agcagcggaagagagagtgatg	380

Search completed: February 13, 2002, 18:47:17
Job time: 18569 sec

Oy	464	gcggccttgaggcgactggaatgtgcgcgttcgagctgtgacgttgcgaagccctgcgtgccccctccctc	523
Db	131	GCAGGCTGGGTGCTGTGAATGGCCGCTGGCTGGCTGTACTGGGCTGGCTGGCCCCCCTCCCTC	190
Oy	544	atgcgtctctcaatgtgtctcatgtaccgcagaaccgcgcgtctctctctgtacctgaaccgaacag	583
Db	191	ATGCGTCTTCATGTGCTTGCTTCATGGCCGAGACCCCGGCGCTTCGTGTGACTTCAGCACAGG	250
Oy	564	cgcgaagagagccat---cgccctgcggtctctctgtgtgggctctccgaagcagaagcttgtgaagac	640
Db	251	CGCCAGGAGGCCCATGTGCCCCTGCGCTGTTCTGTGGGGCTCCGAGCAGCGCTGGGAAGAC	310
Oy	641	cccccaacaggggctgaaacagagctttcaaccttgaccctgcctgcgtgcagccaagccgcgtattac	700
Db	311	CCCCCATTCGGGGCTGAGACAGACGCTTTACCCTGGCGCTCGTCTGGCGACAGCCGCGCATCTAC	370
Oy	701	aagccctcatcatcgcgtgtgtctcccgtatggcgtcttccagagcgtctgcgagggtcaacgcc	760
Db	371	AAGCCCTTCATCATTCGGCGCTCTCCCTGTATGGCTTTCCAGCAAGCTGTGGGGGTTCACAGCCG	430
Oy	761	gtcacgtgtctatgtcagaagaccalcctlttaagaaggccaagtccaaggaagaagagcgtggcc	820
Db	431	GTCATGTTCTATGCGAGAGACCATCTTTGAAAGAGGCCAAGTTCAAAGCAAGCAGCGCTGGCC	490
Oy	821	tccgctgcgtctggtgtgtatcatccacaagctgctgtgtctcaagcgtctgtgcgcgtctcatcatgac	880
Db	491	TGCGTGCCTGCTGGGGTGCATGCAGGTGCTGTTTCAACAGCTGTGGGCGCTGCATCATGTGAC	550
Oy	881	aagagcagggcggaagagcgtgctctccgtgtctgtctgtcgaagtggtgtgtcatatgtgtataagcaaat	940
Db	551	AGAGCAGGGCCGAGAGCTGCTCTCTGTGTGTGACAGGTGTGTGCATAGTGTTTACACCAAGT	610
Oy	941	gacctgcgcgtactacttaagactaacccaaggtgtgcccgtgacaactccctgcacagtgacc	1000
Db	611	GCTTCGGCGCGCTACTTCAAGCTGACCCAGCGGTGGCCCTGGCAACTCTCTGSCACGTGGCC	670
Oy	1001	atctcgagcgctgtctctgc	1020
Db	671	ATCTGGCGCGCTGTCTGTGC	690
RESULT	14		
LOCUS	AIJ38005/c		
DEFINITION	AIJ38005	451 bp	mRNA
VERSION	q35d10.x1 Soares-Pregnant_uterus_NbHnU Homo sapiens cDNA clone	EST	13-FEB-1999
KEYWORDS	AIJ38005	IMAGE:1949971 3,	mRNA sequence.
SOURCE	AIJ38005.1	GI:4074932	EST.
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
JOURNAL	1 (bases 1 to 451)		
COMMENT	NCI-CGAP http://www.ncbi.nlm.nih.gov/cgi/gap.		
FEATURES	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
source	Tumor Gene Index		
	Unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapsb-femail.nih.gov		
	This clone is available royalty-free through LNLN ; contact the		
	IMAGE Consortium (info@image.lnl.gov) for further information.		
	Insert length: 515 Std Error: 0.00		
	Seq primer: -40UP from Gibco.		
	Location/Qualifiers		
	I. 451		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:1949971"		
	/sex="female"		
	/dev_stage="adult"		
	/lab_host="DH10B"		

```

/note="Organ: uterus; Vector: pT73-Pac; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(47) primer [5',
AATCGAAGAATTCGCGCGCCCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by M. Fatima Bonaldo."

```

Query Match	21.2%	Score 384	DB 10	Length 451
Best Local Similarity	99.8%	Pred. No. 1e88		
Matches 434	Conservative	0	Mismatches 1	Indels 0
				Gaps 0
QY 1340	gaacaatcacagcccatctttagagagcgatgacagcaccactcaactagaggatgagcaag	1399		
Db 451	GAACAATCACAGCCCATTTTGGAGGGGGGATGATGACAGCCACTCTACTAGGGGATGAGCAAG	392		
QY 1400	ccctgactccaagcttgggcccaagcccaagcccctgctcgtgcccgaaggagccagaat	1459		
Db 391	CCTGTGACTCCCAAGCTGGGCCCAAGCCCAAGACCCCTCTGCTGCCACAGGGAGCCGAAT	332		
QY 1460	ccagccccccttggagcgttggcttgtagaggttcctcctcttcgtatgtctcctcaagccc	1519		
Db 331	CCAGCCCTCTTGAGCTTGTGCTCGAGGGGTCCCTCTCTGTCATGCTCCCTCCAGGCC	272		
QY 1520	atgaccccgagcctagagagctcgaagctcgaagctcctgcttcagctcctgctcgtcgaaga	1579		
Db 271	ATGACCCGGGCTTAGAGAGCTCAGCTCCTCTCTGTTCCAGCTCTCTCTGCTGTCAGAGA	212		
QY 1580	ctcaggaacaccccttcagagcttgcagacctgcygctcagccctcoactgcgaagactaaag	1639		
Db 211	CTCAGGAACACCTTCGAGACTTTTCAGAGCTCGCGGCACCCCTCCATGCGCAAGACTAAG	152		
QY 1640	cagcggaagaagagtgagctcctcagaagctcttgcctctcgtcctgcagagtgacttttag	1699		
Db 151	CAGCGGAAGAGGAGGTGGGCTCTTAGAGATCTTTGCTCTTGCTGAGAGTGCTTTTGGAG	92		
QY 1700	gttgggtgcttggagcattgcgtcgcctcctcctcacgagcgtgctcttaccggaagaacttc	1759		
Db 91	GTTGGTGTCTGGCATTTGGTGGTGGTCTTCACGGGGCGTGGCTTATCGGAAGAAATTT	32		
QY 1760	gtttggccaataaag 1774			
Db 31	GTTTGCCAAATAAAG 17			
RESULT 15				
AL041348				
LOCUS	AL041348	389 bp	mRNA	EST
DEFINITION	DEFZP434P0816_11.434 (synonym: htes3) Homo sapiens CDNA clone			29-FEB-2000
ACCESSION	DEFZP434P0816.5			mRNA sequence.
VERSION	AL041348			
KEYWORDS	AL041348.1	GI:5420699		
SOURCE	EST.			
ORGANISM	human.			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euleostomi;			
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
JOURNAL	1 (Phases 1 to 389)			
COMMENT	Koehrer,K., Beyrer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.			
	EST (Koehrer, et al.)			
	Unpublished (1999)			
	Contact: Koehrer K			
	MIPS			
	Am Klopferapitz 18a D-82152 Martinsried, Germany			
	This is the 5' sequence of the clone insert			
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer			
	Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;			
	sequenced by BMFZ (Biomedical Research Center at the Charite,			


```

LOCUS       AL565390       752 bp      mRNA           EST       16-FEB-2001
DEFINITION   AL565390 LTI_FL013_FBrn1 Homo sapiens cDNA clone CS0DP005YE02.3
ACCESSION   AL565390
VERSION     AL565390.1 GI:12916718
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 752)
AUTHORS      Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 Evry cedex - France
              Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.
              Location/Qualifiers
                1..752
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="CS0DP005YE02"
                /clone_id="LTI_FL013_FBrn1"
                /dev_stage="pooled tissue from post conception fetuses (20
                week, 24 week and 26 week)"
                /lab_host="DH10B"
                /note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand
                cDNA was primed with a Molt-Oligo(dT) primer. Five prime
                end enriched, double-stranded cDNA was digested with Not I
                and cloned into the Not I and Eco RV sites of the
                pCMVSPORT 6 vector. Library was constructed by Life
                Technologies. Contact : Feng Liang Life Technologies, a
                division of Invitrogen 9800 Medical Center Drive Rockville
                , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
                fliang@lifestech.com URL :
                http://fulllength.invitrogen.com"
BASE COUNT   177 a      206 c      242 g      118 t      9 others
ORIGIN
Query Match      23.3%; Score 423; DB 10; Length 752;
Best Local Similarity 100.0%; Pred. No. 7.6e-99;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1183 caccactgctcgtcgtcttcgtgacaaagagtcagccctccatgaggtctc 1242
    |||||||
Db 548 CACCACCTGGCTCATGGCTTTCGTGACCAAGAGATTTCAGACCCCTCATGAGCTCT 489

QY 1243 caggccctatgagcctctcgtcgtccgtccttcgtcactcctcgtcctc 1302
    |||||||
Db 488 CAGGCCCTATGAGACCTTCGTGCGCTTCGCTTCGATCTTCACGCTTTTCAC 429

QY 1303 ttgtctctgtccctgaaactaaagaaagctctggaacaatcacagcccatl 1362
    |||||||
Db 428 TTTGTTCTGTGTCCTTAATAAGMAAGACTCTGGAAACAATTCACAGCCCATTTGA 369

QY 1363 ggggcgaatgaagcactcactaaggagatgagaagcctctgactccaactggcc 1422
    |||||||
Db 368 GGGCGATGACAGCAGCACTTAGGGGATGGAGCAAGCTTGATGACCAACTGGGCCA 309

QY 1423 agcccaagaccctctgctcccaaggagagcaagaatccacccctcgtgagc 1482
    |||||||
Db 308 AGCCCAAGACCCCTGCGCCCAAGGGAGCCAGACATCCACCCCTTGAGACCTTG 249

QY 1483 gcaaggtccctcctctctctgtcactcctcccaagcccatgacccggag 1542
    |||||||
Db 248 GCAGGGTCCCTCTCTCTGTATGCTCCCTCCACCAATGACCCCGGCTTGAGAG 189

QY 1543 ctgctcctctgtccagctcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1602
    |||||||
Db 188 CTGCTCCTCTGTCCAGCTCCTGCTGCTGCTGCTGAGGACTCAGGAAACACCT 129

```

```

QY 1603 cag 1605
    |||
Db 128 CAG 126

RESULT 11
LOCUS       BG478000       828 bp      mRNA           EST       21-MAR-2001
DEFINITION   602522606p1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4641145 5',
              mRNA sequence.
ACCESSION   BG478000
VERSION     BG478000.1 GI:13410279
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 828)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgabbs-remail.nih.gov
              Tissue Procurement: ATCC/DCTD/DTF
              cDNA Library Preparation: Ling Hong/Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.lnl.gov
              Plate: LNCMI407 row: k column: 02
              High quality sequence stop: 788.
              Location/Qualifiers
                1..828
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4641145"
                /clone_id="NIH_MGC_20"
                /tissue_type="melanotic melanoma"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: skin; Vector: pORF7; Site:1; XhoI; Site:2;
                EcoRI; cDNA made by oligo-dT priming. Directionally
                cloned into EcoRI/XhoI sites using the following 5'
                adaptor: GGCACGAG(G). Size-selected >500bp for average
                insert size 1.8kb. Library constructed by Ling Hong in
                the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT   85 a      309 c      275 g      159 t
ORIGIN
Query Match      21.9%; Score 397; DB 11; Length 828;
Best Local Similarity 99.8%; Pred. No. 2.9e-92;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 ggggcgcgcgcgtctctcgtcgcgcctcgcgcgcgcgcgcgcgcgcgcgcgcgc 65
    |||||||
Db 111 GGGGCGCGCGCTCTCTCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCT 170

QY 166 tgcgctcgtctacagcctcccgccatccctcctcctcctcctcctcctcctc 125
    |||||||
Db 171 TCGCGCTCGGCTACAGCTCCCGCGGCAATCCGACGCTCGACGCGCGCGCGCGCG 230

QY 126 cgcgcctgagcagcgcgcgctcctcgtctcgcgcgcgtcgtcgtcgtcgtc 185
    |||||||
Db 231 CCGCGCTTCACAGACGCCCGCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCTTC 290

QY 186 cgggagagtgctgagcgcgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 245
    |||||||
Db 231 CGGGGAGAGTCTGGCGGCGCTGCGTGGAGACCGCGCGCGCGCGCAAGCTTGC 350

QY 246 tgcgtccgtgcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 305

```


Accession	Sequence	Position
Db	CTGTTCACAGCTGTGGCGGCTTCATCATGACAGACAGGGGGGAGGCGTCCCTGGTGC	249
Oy	ctgtcaagctgtgagtcacatggtgttcaacagagagtccttcggcgccacttcaagctaac	967
Db	TTGTACAGTGTGGTCATGTGTTTCACGACACAGTGCCCTTGCGGCCCTACTTCAAGCTGACG	309
Oy	caggatgagcccttgagcaactccctcgcaagtgccactctggagcgcttgcctctacacagct	1027
Db	CAGGCTGGCCCTGGCAACTCTTCGACAGTGGCATCTCGCGCCCTCTCTTCACACAGCT	369
Oy	gtctatgccagcgttgagccttcgacctgagcgcttgagcagatgacctcttcatgc	1087
Db	GTTTCATGCCACAGCGGGGCTGGCCTGGCTGGCGGAGCATGTGCTCTTCATGCC	429
Oy	gaccttcgattgagcttgaggagccatcccccttgactcttcattctgagagac	1138
Db	GGCTTTTCGGTGAGCTGGGGGCCCATTCCTCCCTGGCTCTTCATGTCAAGATG	480

RESULT	7			
LOCUS	A1971919			
DEFINITION	A1971919	457 bp	mrna	EST 27-OCT-1999
ACCESSION	U03008.1	NCI_GGAE_Ov18	Homo sapiens CDNA clone IMAGE:2531031	3'
VERSION	similar to	U039416	Q39416 INTEGRAL MEMBRANE PROTEIN.	sequence.
KEYWORDS	A1971919			
SOURCE	A1971919.1	GI:5768745		
	EST.			
	human.			

ORGANISM Homo sapiens
Eukaryota: Metazoa: Chordata: Cranialia: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
1 (Phases 1 to 457)
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D.

Email: cgapds-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D., CDNA Library Preparation: M. Bento Soares, Ph.D., CDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbdp/image/image.html
Seq primer: -400P from Gibco
High quality sequence stop: 373.

FEATURES	Location/Qualifiers
source	1. .457

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2531031"
/clone_1ib="NCLCGAP_Ov18"
/tissue_type="fibrotheoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pRT73D-Pac (Pharmacia) with
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - o119o(dT) primer: [
TGTTCACATCTGAAGTGGAGCGCGCCGCCACATTTTGTGTGTGTGT 3'
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the NO
I and Eco RI sites of the modified pRT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Ronaldo. "
BASE COUNT
37 a 174 c 157 g 89 t

```

Query Match	25.2%;	Score 457;	DB 10;	Length 457;
Best Local Similarity	100.0%;	Pred. No. 2.7e-107;		

[illegible]

RESULT	8
LOCUS	A1094911
DEFINITION	473 bp mRNA
	NCI CGAP Brn23 Homo sapiens cDNA clone IMAGE:15687085 3'
	01-OCT-1998

similar to TR:Q39416 Q39416 INTEGRAL MEMBRANE PROTEIN. ;, mRNA
sequence.
A1094911
A1094911.1 GI:3433887
EST.
KEYWORDS
SOURCE human.

ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 473)
AUTHORS	NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute / National Institute of Neurological

JOURNAL
COMMENT
Disorders and Stroke, Brain Tumor
(CGAP/BTGA), Tumor Gene Index
Unpublished (1998)
Contact: Robert Strausberg, Ph.D.

Email: cgapbbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/bdrrp/image/image.html
Insert Length: 1597 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 464.

```

FEATURES
SOURCE
Location/Qualifiers
1..473
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cclone="IMAGE:1687085"

```

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM9706 row: 1 column: 01
High quality sequence stop: 710.

FEATURES

Source

Location/Qualifiers

1..934
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3903048"
/clone_lib="NIH.MGC.70"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site:1; NotI; Site:2; SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.1 kb. Library constructed by Life Technologies."

BASE COUNT 143 a 291 c 287 g 213 t
ORIGIN

Query Match 26.8%; Score 486; DB 11; Length 934;
Best Local Similarity 99.8%; Pred. No. 5.9e-115;
Matches 606; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

667 tcaatggccctgctggcgagcccgagatcacaagcccttcacatcgatgctccct 726
21 tcaatggccctgctggcgagcccgagatcacaagcccttcacatcgatgctccct 80
727 gatgaccttcaagcagcgtctcgagggtcaacgctcatgttctctgagaacacatctt 786
81 gatgaccttcaagcagcgtctcgagggtcaacgctcatgttctctgagaacacatctt 140
787 tgaagagcccaagtcgaagagacagcgtgctcgctcgctgctgctgctgctgctgct 846
141 tgaagagcccaagtcgaagagacagcgtgctcgctcgctgctgctgctgctgctgct 200
847 gctgtcaacgtggtggcggtctctcatatgagagagagagagagagagagagagagag 906
201 gctgtcaacgtggtggcggtctctcatatgagagagagagagagagagagagagagag 260
907 ctgtcagggtgtgtcagtggttcaagcagagtgcttgcggcgcttgcggcgcttgcggcg 966
261 ctgtcagggtgtgtcagtggttcaagcagagtgcttgcggcgcttgcggcgcttgcggcg 320
967 ccagagtgagccctgagcaactcctgcagctgagcattcgagcgtctgtctctgacagc 1026
321 cca-ggtggccctgagcaactcctgcagctgagcattcgagcgtctgtctctgacagc 379
1027 tttgtagcagcagctgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgag 1086
380 tttgtagcagcagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgag 439
1087 cagcgttgcagtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgag 1146
440 cagcgttgcagtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgag 499
1147 gcatgtaagagagtgagcag 1206
500 gcatgtaagagagtgagcag 559
1207 cgtgaccaaagagtgagcag 1266
560 cgtgaccaaagagtgagcag 619
1267 tgcctcc 1273
|||||||

Db 620 TGCCTCC 626

RESULT 6

AM249090

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AM249090 480 bp mRNA EST 07-JAN-2000
2820759.5prime NIH_MGC_7 Homo sapiens cdna clone IMAGE:2820759.5',
mRNA sequence.
AM249090
AM249090.1 GI:6592083
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 480)
NIH-MGC http://mgi.cni.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Other ESTs: 2820759.3prime
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov

Tissue Procurement: DCTD/DRP CDNA Library Preparation: Ling
Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
Project Clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LLNL at:
www.hlo.llnl.gov/btbp/image/image.html Base Calling / Quality
Scores: PHRED from University of Washington Genome Center. Vector
Trimming: cross-match from University of Washington Genome Center
PHRAP suite. Poly-T Identification: patmatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:
http://www.genome.washington.edu
Plate: LHCMS row: A column: 16
High quality sequence stop: 455.

FEATURES

Source

Location/Qualifiers

1..480
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2820759"
/clone_lib="NIH.MGC.75"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pORF7; Site:1; XhoI; Site:2;
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGAC(G). Size-selected >500bp for average
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 68 a 150 c 154 g 108 t
ORIGIN

Query Match 26.0%; Score 471; DB 10; Length 480;
Best Local Similarity 100.0%; Pred. No. 7e-111;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

668 caactgagccctgctggcgagcccgagatcacaagcccttcacatcgatgctccctg 727
10 caactgagccctgctggcgagcccgagatcacaagcccttcacatcgatgctccctg 69
728 atggccttccagcagcgtgctggaggttaacagcgtgctgctatgagagacacatctt 787
70 atggccttccagcagcgtgctggaggttaacagcgtgctgctatgagagacacatctt 129
788 gaagagagcaagtgtaag 847
130 gaagagagcaagtgtaag 189
848 ctgttcaagcgtgtgctgctctcatatgagagagagagagagagagagagagagagagagag 907

962	23	1.3	305	11	BC609663	323594	MA
C 963	23	1.3	305	11	BI282433	UI-R-DD0-	
C 964	23	1.3	306	11	AI511251	UI-R-C3-S	
C 965	23	1.3	306	11	BE905046	BE905046	601496869
C 966	23	1.3	307	11	BE654523	UI-M-A1-I-	
C 967	23	1.3	307	11	BI292442	UI-R-DN0-	
C 968	23	1.3	308	10	AI691538	AI691538	WL33b11.x
C 969	23	1.3	309	10	AI629497	AI629497	486101F12
C 970	23	1.3	310	10	AI266260	qX69f06.x	
C 971	23	1.3	310	11	BE990183	UI-M-B2-I-	
C 972	23	1.3	312	10	BE111438	UI-R-B1-I-	
C 973	23	1.3	312	11	BF397510	UI-R-B5-2	
C 974	23	1.3	312	11	BF405196	UI-R-C1-	
C 975	23	1.3	314	10	AA652037	ns48h04.s	
C 976	23	1.3	315	10	AM002645	wu60f10.x	
C 977	23	1.3	315	11	BC686389	602638314	
C 978	23	1.3	315	11	BF085131	RC3-GN002	
C 979	23	1.3	317	10	AA832372	OC92C02.S	
C 980	23	1.3	318	10	BE449567	EST356326	
C 981	23	1.3	318	11	BI344269	372697	MA
C 982	23	1.3	318	11	BF387765	UI-R-C1-	
C 983	23	1.3	318	11	BF690883	602246889	
C 984	23	1.3	319	10	AI395522	MA003983	
C 985	23	1.3	320	11	AI946870	bs32b08.y	
C 986	23	1.3	320	11	BG668830	DRNBSG08	
C 987	23	1.3	321	10	AA765077	n278e02.s	
C 988	23	1.3	321	10	AI754010	cr18e07.x	
C 989	23	1.3	321	10	AM068980	cr39f09.x	
C 990	23	1.3	321	10	AM294402	UI-H-B1-2	
C 991	23	1.3	322	11	BG107728	602277922	
C 992	23	1.3	324	10	AA998022	UI-R-C0-h	
C 993	23	1.3	325	11	BE394605	BF394605	UI-R-C1-K
C 994	23	1.3	326	10	AI045640	UI-R-C1-K	
C 995	23	1.3	327	10	AI434729	cl10a05.x	
C 996	23	1.3	327	10	AA513166	n129d04.s	
C 997	23	1.3	327	11	BF683393	602139554	
C 998	23	1.3	328	10	AA831291	OC60C03.S	
C 999	23	1.3	328	10	AI136473	UI-R-C2P-	
C1000	23	1.3	328	10	AV531599	AV531599	

ALIGNMENTS

RESULT 1
LOCUS BG749509 949 bp mRNA EST 15-MAY-2001
DEFINITION 602707648P1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:484411 5',
mRNA sequence.
ACCESSION BG749509
VERSION BG749509.1 GI:14060162
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 949)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapds-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM1681 row: P column: 12
High quality sequence stop: 745.

FEATURES
source 1..949
Location/Qualifiers

		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:484411"	
		/clone_11b="NIH_MGC_43"	
		/tissue_type="normal pigmented retinal epithelium"	
		/lab_host="DH10B (phage-resistant)"	
		/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library. !"	
BASE COUNT		160 a	302 c 264 g 223 t
ORIGIN			
Query Match		40.0%; Score 725; DB 11; Length 949;	
Best Local Similarity		100.0%; Pred. No. 9.2e-176;	
Matches 725: Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	971	ggatggccctggcaactcctcgcagctggccatctcggcgcctgtctctgcaagcctgtt	1030
DB	29	GGTGGCCCTGGCAACTCCTCGCACGTCGCGCCATCGCGCCCTGTCTCTGCACAGCCTGTT	88
QY	1031	gatgcagcgtgggctgtgctgtgctgtgcgcgtggcagcatgtgctcttcatgcgcgc	1090
DB	89	GATGCCAGCTGGGCTGGCTGGCTGGCGCCGTCGACGATGTGCTTCATACGCCGCG	148
QY	1091	tttcggtgggtgggtgggtgggtgggtgggtgggtgggtgggtgggtgggtgggtgggt	1150
DB	149	TTTTCGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG	208
QY	1151	gtcaagggcgtggcgaagcaatcgtgctcctcaccacacacacacacacacacacacac	1210
DB	209	GTCACAGCGCGTGGCGACGACGATCTGCTTCACCACTGACTGACTGCTTTCGCTG	268
QY	1211	accgaaggatctgaagcctcctcctcctcctcctcctcctcctcctcctcctcctcctc	1270
DB	269	ACCAAGGATCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	328
QY	1271	tcgctcttcgtcatcttcagtgctccttctccttctccttctccttctccttctccttct	1330
DB	329	TCCGCTTCTGATCTTCAGTGTCTTCCTTTCACCTTTCCTTTCCTTTCCTTTCCTTTC	388
QY	1331	aagactctggaaacaatcaccagccatttgaggggcgaatgaagcactcactagggga	1390
DB	389	AAGACTCTGGAACAATCAACACCCATTGAGGGCGCATGACAGCCACTACTAGGGGA	448
QY	1391	tggagcaagcgtgtactccaagcttgggtgggtgggtgggtgggtgggtgggtgggtgg	1450
DB	449	TGGAGCAAGCGTGTACTCAAGCTGGGCGCAAGCCCAAGCCCTGCTGCCCCAGGGG	508
QY	1451	agccgaatccagcccttggagccttggcaggggtccctccttctcgtcatgctcc	1510
DB	509	AGCCGAATCCAGCCCTTGGAGCCTTGGCTGACAGGCTCCTCCTTCTGTCATCTCTCC	568
QY	1511	ctccagccatgaccccggttagaggctcactgctcctcgttctcagctcctgctgtc	1570
DB	569	CTCCAGCCATGACCCCGGGGTAGGAGCTCATCTGCTCTTCTGCTGCTGCTGCTG	628
QY	1571	ctctgaggaactgaagaacaccttgagcttgcagacctgggttaagccctcctcaggg	1630
DB	629	CTCTGAGGAACCTGAGAACCTTTCAGCTTTCAGACCTGCGGTAGCCCTCCATCGCA	688
QY	1631	aagactaaacgaaggaaggaaggttgggtccttagatcttgcctcctcgttgaaggt	1690
DB	689	AGACTAAACGACGGAAGAGGAGGTGGCTCTAGATCTTGTCTCTGCGCTGAGGCTG	748
QY	1691	ctctt 1695	
DB	749	CTTTT 753	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 13:37:48 ; Search time 2841.05 Seconds
(without alignments)
6857.358 Million cell updates/sec

Title: US-09-516-493-6

Perfect score: 1813

Sequence: 1 aacttcgagccgcgcgtctc.....aaaaaaaaaaattcc 1813

Scoring table: OLIGO_NUC
Gapop 60.0 , Capext 60.0

Searched: 11351937 seqs, 5372889281 residues

Word size: 0

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database:

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: qb_estl:*
11: qb_est2:*
12: qb_hlc:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vtl:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	725	40.0	949	11	BG749509 602707648
2	592	32.7	777	11	BG717034 602689093
3	576	31.8	879	11	BG771736 602720404
4	507	28.0	932	11	BG479842 602527373
5	486	26.8	934	11	BG910478 601501005
6	471	26.0	480	10	AW249090 2820759.5
7	457	25.2	457	10	AI971919 wv30b08.x
8	448	24.7	473	10	AI094911 qal1B03.x
9	430	23.7	442	10	AW051497 w996b11.x
10	423	23.3	752	10	AL565390
11	397	21.9	828	11	BG478000 602522606
12	390	21.5	651	10	AW955452 EST367522
13	386	21.3	830	11	BG700749 602681616
14	384	21.2	451	10	AL338005 q135d10.x
15	380	21.0	389	10	AL041348 DKE2P434P
16	375	20.7	838	11	BG281777 602402292
17	347	19.1	441	10	AL138697 q146e08.x
18	345	19.0	991	11	B1334832 602998959
19	302	16.7	712	11	BG290178 602385265
20	299	16.5	396	10	AI025047 qv6b11.x
21	298	16.4	506	10	BE261639 601149349
22	288	15.9	466	10	AI149830 q132e07.x
23	275	15.2	1014	10	BE735697 601304465
24	273	15.1	421	10	AI205662 q931a06.x
25	267	14.7	425	10	AI290245 qm1a06.x
26	254	14.0	460	10	AW055123 w996b11.x
27	249	13.7	677	11	BG707107 602670234
28	233	12.9	329	10	AA327540 EST30972
29	231	12.7	506	10	AI394026 t911f09.x
30	231	12.7	801	11	B114765 602861314
31	231	12.6	920	11	BF308306 601890370
32	229	12.6	487	10	AA293722 t55b05.s
33	226	12.5	374	10	AA300979 xk10f10.x
34	217	12.0	496	10	AW248655 2820759.3
35	211	11.6	927	11	BF207033 601870114
36	194	10.7	379	10	AA999755 os93e03.s
37	194	10.7	442	10	AA436091 z003a11.r
38	187	10.3	479	11	R11726 yf49g06.r1
39	185	10.2	267	10	BE019192 bp52f08.y
40	185	10.2	531	11	H18721 ym45b08.r1
41	181	10.0	434	11	HO9414 y195h03.r1
42	174	9.6	371	10	AA573501 nk99f03.s
43	163	9.0	325	10	AW392008 QV4-SPT23
44	156	8.6	387	10	BE019085 bd47a06.y
45	152	8.4	327	11	BF742266 RCI-BRN04
46	148	8.2	974	10	AL534131 AL534131
47	139	7.7	812	11	BF246058 601853994
48	134	7.4	479	10	AL602456 DKE2P686M
49	130	7.2	358	10	AA627408 nq49d04.s
50	130	7.2	868	11	BF246849 601854209
51	124	6.8	250	10	AA055061 zf16f09.r
52	118	6.5	408	11	T59518 yb65d09.s1
53	118	6.5	431	11	H18637 ym45b08.s1
54	117	6.5	368	11	T59564 yb65d09.r1
55	115	6.3	303	10	AA435991 z003a11.s
56	109	6.0	181	10	AI094408 oy62c02.s
57	108	6.0	204	11	N47139 y53f05.r1
58	105	5.8	430	10	AI818661 wA89d02.x
59	104	5.7	676	10	BE258822 601117349
60	92	5.1	197	10	AA931975 om89h10.s
61	89	4.9	257	10	AA626007 zu87f05.s
62	89	4.9	433	11	H09355 y195h03.s1
63	89	4.9	471	11	N41051 y53f05.s1
64	89	4.9	474	10	AA024984 ze80c07.s
65	88	4.9	433	11	R17700 yf49g06.s1
66	86	4.7	366	10	AA025081 ze80c07.r
67	82	4.5	911	11	BF311341 601896741
68	80	4.4	389	10	AI970024 w979g10.x
69	79	4.4	250	11	H29433 ym60h06.r1
70	79	4.4	331	10	AI142620 OK35f11.s
71	74	4.1	420	10	AW051558 ym45b08.r1
72	66	3.6	279	11	H46443 y016b08.r1
73	58	3.2	133	10	AI628848 yf17c06.x
74	58	3.2	236	10	AI475029 t6m2c12.x
75	58	3.2	369	10	AI126169 q8b1f11.x
76	50	2.8	595	11	BG079217 H3037C03-
77	50	2.8	833	11	BF120840 601757442
78	41	2.3	319	11	H34451 EST711406 R
79	37	2.0	190	10	AA631985 np66d10.s
80	37	2.0	301	10	AI187942 qe13e06.x
81	36	2.0	430	11	BF468631 ut-M-BH3-
82	36	2.0	436	10	AW762507 ur63f01.y
83	36	2.0	623	10	BE282246 601101202
84	36	2.0	917	11	BF140667 601786917
85	35	1.9	255	10	AI447395 mq63f05.x


```

; TITLE: 2alpha (eIF-2alpha)kinase of rabbit reticulocytes:
; TITLE: Homology to yeast GCN2 protein kinase and human
; TITLE: double-stranded-RNA-dependent
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
; VOLUME: 88
; PAGES: 7729-7733
; DATE: September-1991
; US-08-630-524-1

Query Match          1.3%: Score 23; DB 1; Length 2729;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1787 aatcaaaaaaaaaaaaaaaaaa 1809
      |||||||
Db 2705 AATCAAAAAAAAAAAAAAAAAA 2727

RESULT 14
PCT-US93-08131-1
; Sequence 1, Application PC/TUS9308131
; GENERAL INFORMATION:
; APPLICANT: Massachusetts Institute Of Technology
; TITLE OF INVENTION: DNA Encoding The Heme-Regulated
; TITLE OF INVENTION: Eukaryotic Initiation Factor 2 alpha Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08131
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Padst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: MIT 5807
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2729 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Rabbit
; CELL TYPE: Reticulocytes
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 113..2149
; OTHER INFORMATION: /note= "Expression of HRI mRNA in Human erythroid cells,
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 229..249
; OTHER INFORMATION: /note= "Primer used in the amplification of human HRI cDN
; FEATURE:
; NAME/KEY: misc_feature

```

```

; LOCATION: 543..560
; OTHER INFORMATION: /note= "Primer used in the amplification of human HRI cDN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 448..468
; OTHER INFORMATION: /note= "Primer used in the amplification of human HRI cDN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1009..1031
; OTHER INFORMATION: /note= "Primer used in the amplification of a human HRI c
; PUBLICATION INFORMATION:
; AUTHORS: Chen, Jane J.
; AUTHORS: London, Irving M.
; TITLE: Cloning of the cDNA of the heme-regulated
; TITLE: eukaryotic initiation factor 2alpha
; TITLE: (eIF-2alpha)kinase of rabbit reticulocytes:
; TITLE: Homology to yeast GCN2 protein kinase and human
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
; VOLUME: 88
; PAGES: 7729-7733
; DATE: September-1991
; PCT-US93-08131-1

```

```

Query Match          1.3%: Score 23; DB 5; Length 2729;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1787 aatcaaaaaaaaaaaaaaaaaa 1809
      |||||||
Db 2705 AATCAAAAAAAAAAAAAAAAAA 2727

```

```

RESULT 15
US-09-306-290-37/c
; Sequence 37, Application US/09306290
; Patent No. 6221635
; GENERAL INFORMATION:
; APPLICANT: Rohrer, Giovanni
; TITLE OF INVENTION: METHODS FOR SOLID-PHASE AMPLIFICATION OF DNA TEMPLATE
; TITLE OF INVENTION: (SPADT) USING MULTIBARRAYS
; FILE REFERENCE: 09924-10
; CURRENT APPLICATION NUMBER: US/09/306,290
; CURRENT FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 43
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer p41
; OTHER INFORMATION: RH598
; US-09-306-290-37

```

```

Query Match          1.2%: Score 22; DB 4; Length 43;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1788 aatcaaaaaaaaaaaaaaaaaa 1809
      |||||||
Db 24 AATCAAAAAAAAAAAAAAAAAA 3

```

Search completed: February 13, 2002, 20:05:17
 Job time: 18144 sec

```

? ANTI-SENSE: NO
? FRAGMENT TYPE: N-terminal
? ORIGINAL SOURCE:
? ORGANISM: Rabbit
? CELL TYPE: Reticulocytes
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 113..2149 /note= "Expression of HRI
? OTHER INFORMATION: mRNA in Human erythroid cells, using as the
? OTHER INFORMATION: probe rabbit HRI cDNA from nucleotides 113
? OTHER INFORMATION: to 2149."
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 229..249 /note= "primer used in the
? OTHER INFORMATION: amplification of human HRI cDNA sequence
? OTHER INFORMATION: amplification of human HRI cDNA sequence."
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 543..560 /note= "primer used in the
? OTHER INFORMATION: amplification of human HRI cDNA sequence
? OTHER INFORMATION: amplification of human HRI cDNA sequence."
? OTHER INFORMATION: using the rabbit HRI cDNA sequence."
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 448..468 /note= "primer used in the
? OTHER INFORMATION: amplification of human HRI cDNA sequence
? OTHER INFORMATION: amplification of human HRI cDNA sequence."
? OTHER INFORMATION: using the rabbit HRI cDNA sequence."
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 1009..1031 /note= "primer used in the
? OTHER INFORMATION: amplification of a human HRI cDNA sequence
? OTHER INFORMATION: amplification of a human HRI cDNA sequence."
? OTHER INFORMATION: using the rabbit HRI cDNA sequence."
? PUBLICATION INFORMATION:
? AUTHORS: Chen, Jane J.
? TITLE: Cloning of the cDNA of the heme-
? TITLE: regulated eukaryotic initiation factor
? TITLE: 2alpha (eIF-2alpha)kinase of rabbit reticulocytes:
? TITLE: Homology to yeast GCN2 protein kinase and human
? TITLE: double-stranded-RNA-dependent
? JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
? VOLUME: 88
? PAGES: 7729-7733
? DATE: September-1991
US-07-938-782A-1

Query Match 1.3%; Score 23; DB 1; Length 2729;
Best local Similarity 100.0%; Pred. No. 1.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Cy 1787 aaatcaaaaaaaaaaaaaaa 1809
Db 2705 AAATCAAAAAAAAAAAAAAA 2727

RESULT 13
US-08-630-524-1
Sequence 1, Application US/08630524
Patent No. 5650930
GENERAL INFORMATION:
APPLICANT: Chen, Jane J.
TITLE OF INVENTION: DNA Encoding the Heme-Regulated
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: Patrea L. Pabst
STREET: 2800 One Atlantic Center

```

```

STREET: 1201 West Peachtree Street
City: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,524
FILING DATE: 10-APR-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/938,782
FILING DATE: 31-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patricia L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: W0101C1P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)873-8794
TELEFAX: (404)873-8795
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2729 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Rabbit
CELL TYPE: Reticulocytes
FEATURE:
NAME/KEY: misc_feature
LOCATION: 113..2149
OTHER INFORMATION: /note= "Expression of HRI
OTHER INFORMATION: mRNA in Human erythroid cells, using as the
OTHER INFORMATION: probe rabbit HRI cDNA from nucleotides 113
OTHER INFORMATION: to 2149."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 229..249
OTHER INFORMATION: /note= "Primer used in the
OTHER INFORMATION: amplification of human HRI cDNA sequence
OTHER INFORMATION: using the rabbit HRI cDNA sequence."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 543..560
OTHER INFORMATION: /note= "Primer used in the
OTHER INFORMATION: amplification of human HRI cDNA sequence
OTHER INFORMATION: using the rabbit HRI cDNA sequence."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 448..468
OTHER INFORMATION: /note= "Primer used in the
OTHER INFORMATION: amplification of human HRI cDNA sequence
OTHER INFORMATION: using the rabbit HRI cDNA sequence."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1009..1031
OTHER INFORMATION: /note= "Primer used in the
OTHER INFORMATION: amplification of a human HRI cDNA sequence
OTHER INFORMATION: using the rabbit HRI cDNA sequence."
PUBLICATION INFORMATION:
AUTHORS: Chen, Jane J.
AUTHORS: London, Irving M.
TITLE: Cloning of the cDNA of the heme-
TITLE: regulated eukaryotic initiation factor

```

TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-416-050A-7

Query Match 1.3%; Score 23; DB 4; Length 1737;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1787 aatcaaaaaaaaaaaaaaa 1809
|||||
Db 1702 aatcaaaaaaaaaaaaaaa 1724

RESULT 9
US-09-664-800-7
Sequence 7, Application US/09664800
Patent No. 6218527
GENERAL INFORMATION:
APPLICANT: KIM, Soo Young
TITLE OF INVENTION: Abscisic Acid Responsive Element -Binding Transcription Factor

FILE REFERENCE: 1942/42
CURRENT APPLICATION NUMBER: US/09/664,800
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 09/416,050
PRIOR FILING DATE: 1999-10-12
NUMBER OF SEQ ID NOS: 83
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 1737
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-664-800-7

Query Match 1.3%; Score 23; DB 4; Length 1737;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1787 aatcaaaaaaaaaaaaaaa 1809
|||||
Db 1702 aatcaaaaaaaaaaaaaaa 1724

RESULT 10
US-09-665-309-7
Sequence 7, Application US/09665309
Patent No. 6232461
GENERAL INFORMATION:
APPLICANT: KIM, Soo Young
TITLE OF INVENTION: Abscisic Acid Responsive Element -Binding Transcription Factor
FILE REFERENCE: 1942/42
CURRENT APPLICATION NUMBER: US/09/665,309
CURRENT FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 09/416,050
PRIOR FILING DATE: 1999-10-12
NUMBER OF SEQ ID NOS: 83
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 1737
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-665-309-7

Query Match 1.3%; Score 23; DB 4; Length 1737;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1787 aatcaaaaaaaaaaaaaaa 1809
|||||
Db 1702 aatcaaaaaaaaaaaaaaa 1724

RESULT 11
US-09-661-569-7
Sequence 7, Application US/09661569
Patent No. 6245905
GENERAL INFORMATION:
APPLICANT: KIM, Soo Young
TITLE OF INVENTION: Abscisic Acid Responsive Element -Binding Transcription Fac
FILE REFERENCE: 1942/42
CURRENT APPLICATION NUMBER: US/09/661,569
CURRENT FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 09/416,050
PRIOR FILING DATE: 1999-10-12
NUMBER OF SEQ ID NOS: 83
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 1737
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-661-569-7

Query Match 1.3%; Score 23; DB 4; Length 1737;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1787 aatcaaaaaaaaaaaaaaa 1809
|||||
Db 1702 aatcaaaaaaaaaaaaaaa 1724

RESULT 12
US-07-938-782A-1
Sequence 1, Application US/07938782A
Patent No. 5525513
GENERAL INFORMATION:
APPLICANT: Chen, Jane J.
TITLE OF INVENTION: DNA Encoding the Heme-Regulated
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/938,782A
FILING DATE: 31-AUG-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: WU101CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)873-8794
TELEFAX: (404)873-8795
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2729 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO

MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 23..1831
US-08-168-091A-1

Query Match 1.3%; Score 24; DB 1; Length 2351;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1786 aaatcaaaaaaaaaaaaaa 1809
|||||
DB 2297 AAATCAAAAAAAAAAAAAA 2320

RESULT 6
US-08-697-954-1
Sequence 1, Application US/08697954
Patent No. 6284535

GENERAL INFORMATION:

APPLICANT: Role, Lorna M.
TITLE OF INVENTION: SPICE VARIANTS OF THE HEREGULIN GENE, NARIA, AND

TITLE OF INVENTION: USES THEREOF

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/697,954

FILING DATE:

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 46839-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-278-0400

TELEFAX: 212-391-0526

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3212 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-697-954-1

Query Match 1.3%; Score 24; DB 4; Length 3212;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1786 aaatcaaaaaaaaaaaaaa 1809
|||||
DB 3131 AAATCAAAAAAAAAAAAAA 3154

RESULT 7
US-08-076-011-1

Sequence 1, Application US/08076011

Patent No. 5521069

GENERAL INFORMATION:

APPLICANT: ONDA, Haruo

APPLICANT: KIMURA, Chiharu

APPLICANT: OHKUBO, Shoichi
TITLE OF INVENTION: NOVEL DNA AND USE THEREOF
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:

ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN,
ADDRESSEE: ROBERTS & CUSHMAN

STREET: 130 Water Street

CITY: Boston

STATE: Massachusetts

COUNTRY: US

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/076,011

FILING DATE: 11-JUN-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/047,246

FILING DATE: 13-APR-1993

NAME:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/741,676

FILING DATE: 07-AUG-1991

NAME:

ATTORNEY/AGENT INFORMATION:

NAME: RESNICK, DAVID S

REGISTRATION NUMBER: 34235

REFERENCE/DOCKET NUMBER: 41155-CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)523-3400

TELEFAX: (617)523-6440

TELEX: 200291 STRE UR

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 17041 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: join(7540..7650, 9814..9945, 10421..10519,
LOCATION: 11602..11787)

US-08-076-011-1

Query Match 1.3%; Score 24; DB 1; Length 17041;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1790 tcaaaaaaaaaaaaaaaaaattcc 1813
|||||
DB 220 TCAAAAAAAAAAAAAAAAAATTCC 243

RESULT 8
US-09-416-050A-7

Sequence 7, Application US/09416050A

Patent No. 6194559

GENERAL INFORMATION:

APPLICANT: KIM, Soo Young

TITLE OF INVENTION: Abscisic Acid Responsive Element - Binding Transcription Facto

FILE REFERENCE: 1942/42

CURRENT APPLICATION NUMBER: US/09/416,050A

CURRENT FILING DATE: 1999-10-12

NUMBER OF SEQ ID NOS: 83

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 1737

APPLICANT: FASEL, NICOLAS JOSEPH
APPLICANT: REYMOND, CHRISTOPHE DOMINIQUE
TITLE OF INVENTION: DICTYOSTELID EXPRESSION VECTOR AND
TITLE OF INVENTION: METHOD FOR EXPRESSING A DESIRED PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: THE WEBB LAW FIRM
STREET: 700 KOPPERS BUILDING, 436 SEVENTH AVENUE
CITY: PITTSBURGH
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 15219-1818
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISK
COMPUTER: Midwest Micro 486-50
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,405A
FILING DATE: 26-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,273
FILING DATE: 15-JAN-1993
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 731
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: UNKNOWN
US-08-451-405A-2

Query Match 1.3%; Score 24; DB 1; Length 731;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1786 aaatcaaaaaaaaaaaaaa 1809
|||||
DB 292 AAATCAAAAAAAAAAAAAA 315
RESULT 4
US-08-568-147B-1
Sequence 1, Application US/08568147B
Patent No. 5783422
GENERAL INFORMATION:
APPLICANT: Suminami, Yoshinori
APPLICANT: Kato, Hiroshi
APPLICANT: Sekiguchi, Kiyoshi
APPLICANT: Takeda, Katsumichi
TITLE OF INVENTION: DNA FRAGMENT CODING FOR SQUAMOUS CELL
TITLE OF INVENTION: CARCINOMA-ASSOCIATED ANTIGEN
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/568,147B
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 800,952
FILING DATE: 02-DEC-1991

ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8425
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1711 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 62..1234
US-08-568-147B-1

Query Match 1.3%; Score 24; DB 1; Length 1711;
Best Local Similarity 100.0%; Pred. No. 0.7;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1786 aaatcaaaaaaaaaaaaaa 1809
|||||
DB 1688 AAATCAAAAAAAAAAAAAA 1711

RESULT 5
US-08-168-091A-1
Sequence 1, Application US/08168091A
Patent No. 565862
GENERAL INFORMATION:
APPLICANT: Fischbach, Gerald.
APPLICANT: Falls, Douglas R.
APPLICANT: Rosen, Kenneth M.
APPLICANT: Cortas, Gabriel
TITLE OF INVENTION: Neurotrophic Factor
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE AND COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/168,091A
FILING DATE: 15-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/953,742
FILING DATE: 29-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: HMT-002CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2351 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

```
C 977 19 1.0 42 5 PCT-US94-02252A-11 Sequence 11, Appl
C 978 19 1.0 43 4 US-08-991-840A-13 Sequence 13, Appl
C 979 19 1.0 46 1 US-07-901-707-21 Sequence 21, Appl
C 980 19 1.0 46 1 US-07-988-430-21 Sequence 21, Appl
C 981 19 1.0 46 1 US-08-425-336-21 Sequence 21, Appl
C 982 19 1.0 46 1 US-08-232-463-40 Sequence 40, Appl
C 983 19 1.0 46 1 US-08-488-113B-21 Sequence 21, Appl
C 984 19 1.0 46 1 US-08-477-484B-21 Sequence 21, Appl
C 985 19 1.0 46 2 US-08-646-360-21 Sequence 21, Appl
C 986 19 1.0 46 3 US-09-322-692-1 Sequence 1, Appl
C 987 19 1.0 46 3 US-08-839-765-21 Sequence 21, Appl
C 988 19 1.0 46 3 US-09-136-389-21 Sequence 21, Appl
C 989 19 1.0 46 5 PCT-US92-09487-21 Sequence 21, Appl
C 990 19 1.0 46 5 PCT-US94-05407-3 Sequence 3, Appl
C 991 19 1.0 46 5 PCT-US94-05407-10 Sequence 10, Appl
C 992 19 1.0 48 1 US-08-380-438-6 Sequence 6, Appl
C 993 19 1.0 48 2 US-08-975-902-40 Sequence 40, Appl
C 994 19 1.0 48 2 US-08-906-713-6 Sequence 6, Appl
C 995 19 1.0 48 3 US-09-251-565-40 Sequence 40, Appl
C 996 19 1.0 49 4 US-08-906-156A-88 Sequence 88, Appl
C 997 19 1.0 49 5 PCT-US96-07709-39 Sequence 39, Appl
C 998 19 1.0 49 5 PCT-US96-07709-40 Sequence 40, Appl
C 999 19 1.0 49 5 PCT-US96-09848-19 Sequence 19, Appl
1000 19 1.0 49 5 PCT-US96-09848-20 Sequence 20, Appl
```

ALIGNMENTS

```
RESULT 1
US-08-574-959A-3
; Sequence 3, Application US/08574959A
; Patent No. 5962224
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
; APPLICANT: and Jack L. Strominger
; TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/574,959A
; FILING DATE: 19-DEC-95
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1977 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1260
; US-08-574-959A-3
```

```
Query Match 1.4%; Score 26; DB 2; Length 1977;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1788 aatcaaaaaaaaaaaaaattcc 1813
Db 1939 AATCAAAAAAAAAAAAAAATTCC 1964
```

```
RESULT 2
US-09-357-014-3
; Sequence 3, Application US/09357014
; Patent No. 6291645
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
; APPLICANT: and Jack L. Strominger
; TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/357,014
; FILING DATE: 19-Jul-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/574,959
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1977 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1260
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-357-014-3

Query Match 1.4%; Score 26; DB 4; Length 1977;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1788 aatcaaaaaaaaaaaaaattcc 1813
Db 1939 AATCAAAAAAAAAAAAAAATTCC 1964

RESULT 3
US-08-451-405A-2
; Sequence 2, Application US/08451405A
; Patent No. 5736358
; GENERAL INFORMATION:
```


CC transgenic animals and the TANGO polypeptides themselves. Partial
 CC TANGO-128, 140, 197, 212, 213, 224, 239 sequences are useful in
 CC forensic biology, for diagnostic assays, prognostic assays,
 CC pharmacogenomics and for monitoring clinical trials. TANGO
 CC polypeptides are suitable for both prophylactic and therapeutic
 CC methods for treating a subject at risk of a disorder or having a
 CC disorder associated with aberrant TANGO expression. A wide range
 CC of cellular disorders can be treated.
 XX
 SO Sequence 3413 BP; 996 A; 734 C; 747 G; 936 T; 0 other:

Query Match 1.4%; Score 25; DB 21; Length 3413;
 Best Local Similarity 100.0%; Pred. No. 3.3;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1785 gaaatcaaaaaaaaaaaaaa 1809
 ||||||||||||||||||||
 Db 3336 gaaatcaaaaaaaaaaaaaa 3360

RESULT 15
 AAF21711
 ID AAF21711 standard; DNA: 249 BP.
 AC AAF21711;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Human breast and ovarian cancer associated antigen gene SEQ ID 98.
 XX
 KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW neoplastic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antihypertensive; antitumor; antineoplastic; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease; ds.
 XX
 OS Homo sapiens.
 XX
 PN W0200055173-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05881.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI: 2000-611515/58.
 DR P-PSDB: AAB58808.
 XX
 XX
 PT New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention,
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular
 PT disorders and neurological diseases -
 XX
 PS Claim 1; Page 555; 1299pp; English.
 XX
 CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;
 CC neoplastic; neuroprotective; antiviral; antiallergic; hepatotropic;
 CC antidiabetic; antiinflammatory; antitumor; antineoplastic; anticonvulsant;

CC antibacterial; antifungal; antiparasitic and cardiant activity. The
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and antagonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemias; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases.
 XX
 SO Sequence 249 BP; 93 A; 51 C; 48 G; 55 T; 2 other:

Query Match 1.3%; Score 24; DB 21; Length 249;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1786 aaatcaaaaaaaaaaaaaa 1809
 ||||||||||||||||||||
 Db 200 aaatcaaaaaaaaaaaaaa 223

Search completed: February 13, 2002, 20:10:13
 Job time: 11465 sec

CC therapeutic, prognostic and diagnostic tools for proliferative and
CC developmental disorders, e.g. neoplasia.
XX
SQ Sequence 1222 BP; 406 A; 206 C; 221 G; 387 T; 2 other;

Query Match 1.4%; Score 25; DB 20; Length 1222;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1785 gaaatcaaaaaaaaaaaaaa 1809
Db 249 gaaatcaaaaaaaaaaaaaa 273

RESULT 13
AAA47459
ID AAA47459 standard; cDNA: 3413 BP.

XX AAA47459;

XX 20-OCT-2000 (first entry)

XX Human TANGO 239 coding sequence.

DE TANGO: 128: 140; 197; 212; 213; 224; 239; modulating agent; asthma;
KW graft versus-host diseases; rheumatoid arthritis; psoriasis;
KW inflammatory bowel disease; septic shock; ulcerative colitis;
KW Crohn's disease; chronic myelogenous leukemia; cancer; liver
KW disease; Hodgkin's disease; osteoarthritis; Lyme's disease;
KW cachexia; autoimmune disease; myasthenia gravis; autoimmune
KW systemic lupus erythematosus; transgenic animal; diagnosis;
KW prognosis; prophylactic; therapeutic; human; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 344..1993

FT /tag= a

FT /product= TANGO 239

XX WO200039284-A1.

XX 06-JUL-2000.

XX 23-DEC-1999; 99WO-US31025.

XX 30-DEC-1998; 98US-0223546.

XX (MILL-) MILLENNIUM PHARM INC.

XX Holtzman DA;

XX WPI: 2000-465743/40.

DR P-PSDB: AAB01426.

PT Novel nucleic acid sequences encoding TANGO-128, 140, 197, 212, 213,
PT 224 and 239 polypeptides useful for the treatment of asthma, rheumatoid
PT arthritis, psoriasis and autoimmune diseases

XX Claim 1: Fig 8; 209pp; English.

XX Nucleic acids encoding TANGO polypeptides are useful as modulating
CC agents for regulating cellular processes like asthma, graft
CC versus-host diseases, rheumatoid arthritis, psoriasis, inflammatory
CC bowel disease, septic shock, ulcerative colitis, Crohn's disease,
CC chronic myelogenous leukemia, cancer, liver disease, Hodgkin's
CC disease, osteoarthritis, Lyme's disease, cachexia and autoimmune
CC diseases e.g. myasthenia gravis, autoimmune diabetes and systemic
CC lupus erythematosus. The nucleic acids are also useful for producing
CC transgenic animals and the TANGO polypeptides themselves. Partial
CC TANGO-128, 140, 197, 212, 213, 224, 239 sequences are useful in
CC forensic biology, for diagnostic assays, prognostic assays,

CC pharmacogenomics and for monitoring clinical trials. TANGO
CC polypeptides are suitable for both prophylactic and therapeutic
CC methods for treating a subject at risk of a disorder or having a
CC disorder associated with aberrant TANGO expression. A wide range
CC of cellular disorders can be treated.

SQ Sequence 3413 BP; 992 A; 733 C; 752 G; 936 T; 0 other;

Query Match 1.4%; Score 25; DB 21; Length 3413;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1785 gaaatcaaaaaaaaaaaaaa 1809
Db 3335 gaaatcaaaaaaaaaaaaaa 3359

RESULT 14
AAA47483
ID AAA47483 standard; cDNA: 3413 BP.

XX AAA47483;

XX 20-OCT-2000 (first entry)

XX Human TANGO 239 coding sequence (form 2).

DE TANGO: 128: 140; 197; 212; 213; 224; 239; modulating agent; asthma;
KW graft versus-host diseases; rheumatoid arthritis; psoriasis;
KW inflammatory bowel disease; septic shock; ulcerative colitis;
KW Crohn's disease; chronic myelogenous leukemia; cancer; liver
KW disease; Hodgkin's disease; osteoarthritis; Lyme's disease;
KW cachexia; autoimmune disease; myasthenia gravis; autoimmune diabetes;
KW systemic lupus erythematosus; transgenic animal; diagnosis;
KW prognosis; prophylactic; therapeutic; human; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 344..2404

FT /tag= a

FT /product= TANGO 239 (form 2)

XX WO200039284-A1.

XX 06-JUL-2000.

XX 23-DEC-1999; 99WO-US31025.

XX 30-DEC-1998; 98US-0223546.

XX (MILL-) MILLENNIUM PHARM INC.

XX Holtzman DA;

XX WPI: 2000-465743/40.

DR P-PSDB: AAB01432.

PT Novel nucleic acid sequences encoding TANGO-128, 140, 197, 212, 213,
PT 224 and 239 polypeptides useful for the treatment of asthma, rheumatoid
PT arthritis, psoriasis and autoimmune diseases

XX Claim 1: Fig 31; 209pp; English.

XX Nucleic acids encoding TANGO polypeptides are useful as modulating
CC agents for regulating cellular processes like asthma, graft
CC versus-host diseases, rheumatoid arthritis, psoriasis, inflammatory
CC bowel disease, septic shock, ulcerative colitis, Crohn's disease,
CC chronic myelogenous leukemia, cancer, liver disease, Hodgkin's
CC disease, osteoarthritis, Lyme's disease, cachexia and autoimmune
CC diseases e.g. myasthenia gravis, autoimmune diabetes and systemic
CC lupus erythematosus. The nucleic acids are also useful for producing

CC in a first filamentous fungal (FF) cell relative to expression of the
 CC same genes in one or more second filamentous fungal cells. Monitoring
 CC the global expression of genes from FF cells allows the production
 CC potential of the microorganisms to be improved. New genes may be
 CC discovered, possible functions of unknown open reading frames can be
 CC identified and gene copy number variation and stability can be
 CC monitored. The expression of genes can be used to study how FF cells
 CC adapt to changes in culture conditions, environmental stress, spore
 CC morphogenesis, recombination, metabolic or catabolic pathway
 CC engineering. Using ESTs provides several advantages over genomic or
 CC random cDNA clones including elimination of redundancy as one spot on an
 CC array equals one gene or open reading frame, and organisation of the
 CC microarrays based on function of the gene products to facilitate
 CC analysis of the results. AAF07478 to AAF11247 represents ESTs from
 CC *Fusarium venenatum*; AAF11248 to AAF11853 represents ESTs from
 CC *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from *Aspergillus*
 CC *niger*; AAF14879 to AAF15337 represents ESTs from *Trichoderma reesei*, which are
 CC all specifically claimed in the present invention.

XX
 SQ Sequence 1008 BP; 287 A; 216 C; 276 G; 229 T; 0 other;

Query Match 1.4%; Score 25; DB 21; Length 1008;
 Best Local Similarity 100.0%; Pred. No. 4;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1789 atcaaaaaaaaaaaaaaatcc 1813
 Db 26 ATCAAAAAAAAAAAAAAATTCC 2

RESULT 11

AAFI2939 ID AAF12939 standard; cDNA: 1149 BP.

XX AAF12939;

XX 13-MAR-2001 (first entry)

XX *Aspergillus oryzae* EST SEQ ID NO:5462.

XX Multiple gene expression; filamentous fungal cell; ESN;
 KW expressed sequence tag; *Fusarium venenatum*; *Aspergillus niger*;
 KW *Aspergillus oryzae*; *Trichoderma reesei*; identification; recombination;
 KW culture condition; environmental stress; spore morphogenesis;
 KW metabolic pathway engineering; catabolic pathway engineering; ss.

XX *Aspergillus oryzae*.

XX WO200056762-A2.

XX 28-SEP-2000.

XX 22-MAR-2000; 2000WO-US07781.

XX 22-MAR-1999; 99US-0273623.

XX (NOVO) NOVO NORDISK BIOTECH INC.

XX (NOVO) NOVO NORDISK AS.

XX Berka RM, Rey MM, Shuster JR, Kaupinen S, Clausen IG, Olsen PB;

XX WPI; 2000-594572/56.

XX Monitoring differential expression of genes in filamentous fungal cells
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a
 PT substrate of expressed sequence tags -

XX Claim 88; Page 2271-2272; 3161pp; English.

XX The present invention describes a method for monitoring differential
 CC expression of genes in a first filamentous fungal (FF) cell relative to
 CC expression of the same genes in one or more second filamentous fungal

CC cells. The method uses fluorescence-labeled nucleic acids isolated from
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
 CC are used in the methods for monitoring differential expression of genes
 CC in a first filamentous fungal (FF) cell relative to expression of the
 CC same genes in one or more second filamentous fungal cells. Monitoring
 CC the global expression of genes from FF cells allows the production
 CC potential of the microorganisms to be improved. New genes may be
 CC discovered, possible functions of unknown open reading frames can be
 CC identified and gene copy number variation and stability can be
 CC monitored. The expression of genes can be used to study how FF cells
 CC adapt to changes in culture conditions, environmental stress, spore
 CC morphogenesis, recombination, metabolic or catabolic pathway
 CC engineering. Using ESTs provides several advantages over genomic or
 CC random cDNA clones including elimination of redundancy as one spot on an
 CC array equals one gene or open reading frame, and organisation of the
 CC microarrays based on function of the gene products to facilitate
 CC analysis of the results. AAF07478 to AAF11247 represents ESTs from
 CC *Fusarium venenatum*; AAF11248 to AAF11853 represents ESTs from *Aspergillus*
 CC *niger*; AAF11854 to AAF14878 represents ESTs from *Aspergillus oryzae*; and
 CC AAF14879 to AAF15337 represents ESTs from *Trichoderma reesei*, which are
 CC all specifically claimed in the present invention.

XX
 SQ Sequence 1149 BP; 220 A; 389 C; 230 G; 310 T; 0 other;

Query Match 1.4%; Score 25; DB 21; Length 1149;
 Best Local Similarity 100.0%; Pred. No. 3; 9;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1789 atcaaaaaaaaaaaaaaatcc 1813
 Db 1124 atcaaaaaaaaaaaaaaatcc 1148

RESULT 12

AAK83994 ID AAX83994 standard; cDNA: 1222 BP.

XX AAX83994;

XX 09-SEP-1999 (first entry)

XX Human EPC1 encoding cDNA.

XX EPC1; EPC2; enhancer of polycomb; tumour suppressor; diagnosis;
 KW proliferative disorder; developmental disorder; neoplasia; ss.

XX *Homo sapiens*.

XX WO9933986-A1.

XX 08-JUL-1999.

XX 28-DEC-1998; 98WO-US27667.

XX 30-DEC-1997; 97US-0068957.

XX (CHIR) CHIRON CORP.

XX Randazzo F;

XX WPI; 1999-419108/35.

XX Mammalian tumour suppressor, termed enhancer of polycomb, useful for
 PT therapy, prognosis and diagnosis of proliferative and developmental
 PT disorders

XX Claim 15; Page 50; 55pp; English.

XX The present sequence encodes a human tumour suppressor, designated
 CC enhancer of polycomb (Epc). Human Epc genes can be used to identify a
 CC p10-11 region of human chromosome 10 and a q13.3 region of human
 CC chromosome 22. The Epc genes and expression products are useful as

CC suppressor genes. (1), and polypeptides encoded by them, are useful as
CC targets for diagnosis or therapy and in screening to determine the
CC effects of an active compound (potential pharmaceutical) on a cell line,
CC particularly for diagnosis and treatment of tumors, especially by
CC modulating expression of (1) (by gene therapy, antisense RNA or ribozyme
CC methods) or by modulating the amount and/or location of (1)-encoded
CC polypeptides (by administration of the polypeptide or its activator,
CC antibody (optionally as a conjugate) or inhibitor). The method allows
CC identification of many Class II tumour suppressor genes (i.e. genes that
CC are not primary targets for tumour-initiating mutations).
CC AAH81492-AAH82376 represent the human and rat derived nucleic acid
CC fragments described in the method of the invention.
CC
XX
SQ Sequence 251 BP; 76 A; 31 C; 28 G; 116 T; 0 other;

Query Match 1.4%; Score 25; DB 22; Length 251;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1785 gaatacaaaaaaaaaaaaaa 1809
|||||
Db 53 GAAATCAAAAAAAAAAAAAA 29

RESULT 9
AAf08921/c
ID AAF08921 standard; cDNA: 567 BP.
AC AAF08921;
XX
DT 13-MAR-2001 (first entry)
XX
DE Fusarium venenatum EST SEQ ID NO:1444.
XX
XX Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS Fusarium venenatum.
PN WO200056762-A2.
XX
PD 28-SEP-2000.
XX
PF 22-MAR-2000; 2000WO-US07781.
XX
PR 22-MAR-1999; 99US-0273623.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
PA (NOVO) NOVO NORDISK AS.
XX
PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
XX
XX WPI: 2000-594572/56.
XX
PT Monitoring differential expression of genes in filamentous fungal cells
PT uses fluorescence-labeled nucleic acids isolated from the cells and a
PT substrate of expressed sequence tags -
XX
PS Claim 86: Page 927; 3161pp: English.
XX
XX The present invention describes a method for monitoring differential
CC expression of genes in a first filamentous fungal (FF) cell relative to
CC expression of the same genes in one or more second filamentous fungal
CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes
CC in a first filamentous fungal (FF) cell relative to expression of the
CC same genes in one or more second filamentous fungal cells. Monitoring
CC the global expression of genes from FF cells allows the production

CC potential of the microorganisms to be improved. New genes may be
CC discovered, possible functions of unknown open reading frames can be
CC identified and gene copy number variation and stability can be
CC monitored. The expression of genes can be used to study how FF cells
CC adapt to changes in culture conditions, environmental stress, spore
CC morphogenesis, recombination, metabolic or catabolic pathway
CC engineering. Using ESTs provides several advantages over genomic or
CC random cDNA clones including elimination of redundancy as one spot on an
CC array equals one gene or open reading frame, and organisation of the
CC microarrays based on function of the gene products to facilitate
CC analysis of the results. AAF07478 to AAF11247 represents ESTs from
CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
CC niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and
CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
CC all specifically claimed in the present invention.
CC
XX
SQ Sequence 567 BP; 136 A; 127 C; 167 G; 136 T; 1 other;

Query Match 1.4%; Score 25; DB 21; Length 567;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1789 atcaaaaaaaaaaaaaaatcc 1813
|||||
Db 31 ATCAAAAAAAAAAAAAAATTC 7

RESULT 10
AAf12934/c
ID AAF12934 standard; cDNA: 1008 BP.
AC AAF12934;
XX
DT 13-MAR-2001 (first entry)
XX
DE Aspergillus oryzae EST SEQ ID NO:5457.
XX
XX Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS Aspergillus oryzae.
PN WO200056762-A2.
XX
PD 28-SEP-2000.
XX
PF 22-MAR-2000; 2000WO-US07781.
XX
PR 22-MAR-1999; 99US-0273623.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
PA (NOVO) NOVO NORDISK AS.
XX
PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
XX
XX WPI: 2000-594572/56.
XX
PT Monitoring differential expression of genes in filamentous fungal cells
PT uses fluorescence-labeled nucleic acids isolated from the cells and a
PT substrate of expressed sequence tags -
XX
PS Claim 88: Page 2269-2270; 3161pp: English.
XX
XX The present invention describes a method for monitoring differential
CC expression of genes in a first filamentous fungal (FF) cell relative to
CC expression of the same genes in one or more second filamentous fungal
CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes

PA (UYLA-) UNTV LAUSANNE.
 XX
 PI Thorens B, Idberson M, Uldry M;
 XX
 DR WPI: 2001-112615/12.
 DR P-PSDB: AAB66933.
 XX
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -
 XX
 PS Claim 3: Page 71-73; 124pp; English.
 XX
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is the coding sequence
 CC for rat GLUTX1.
 CC
 SQ Sequence 2087 BP: 355 A; 673 C; 580 G; 479 T; 0 other:
 Query Match 2.3%; Score 41; DB 22; Length 2087;
 Best Local Similarity 100.0%; Pred. No. 7.2e-06;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1109 cccatccctgcctcctcatgcaagatccctctgca 1149
 Db 1203 cccatccctgcctcctcatgcaagatccctctgca 1243
 RESULT 7
 AAT89344
 ID AAT89344 standard; cDNA; 1977 BP.
 XX
 AC AAT89344;
 XX
 DT 11-MAR-1998 (first entry)
 XX
 DE Human p62 cDNA #2.
 XX
 KW p62; cytoplasmic; T cell; B cell; development; activation; modulation;
 KW cellular response; cell proliferation; autoimmune disease; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..1260
 FT /*tag= a
 FT /product= p62
 FT /note= "partial CDS"
 XX
 PN W09722255-A1.
 XX
 PD 26-JUN-1997.
 XX
 PF 11-DEC-1996; 96MO-US19944.
 XX
 PR 19-DEC-1995; 95US-0574959.
 XX
 PA (DAND) DANA FARBER CANCER INST INC.
 XX
 PI Young I, Shin J, Strominger JL, Vadamudi RK;
 XX
 DR WPI: 1997-341351/31.
 DR P-PSDB: AAM31183.
 XX
 CC cDNA encoding p62 and p160 and corresponding proteins - used in the
 PT treatment of autoimmune disease and for T and B cell proliferation,
 PT e.g. for treatment of tumours

XX
 PS Claim 4: Fig 3; 175pp; English.
 XX
 CC This cDNA sequence encodes a novel p62 cytoplasmic polypeptide which has
 CC 77.5 per cent homology to the nucleic acid sequence of p62 represented
 CC in AAT89343. This cytoplasmic polypeptide is expressed in B cells and
 CC other cells of haematopoietic origin e.g. T cells. p62 is capable of
 CC modulating T or B cell development and/or T or B cell activation e.g. by
 CC modulation of Lck activity. It is also capable of modulating degradation
 CC of cellular proteins e.g. cell cycle regulatory proteins stimulating
 CC expression of cell cycle dependent kinase inhibitors and arresting cell
 CC cycle progression at specific boundaries to thereby modulate cell
 CC proliferation. As p62 acts to boost B cell response and may be used to
 CC treat disorders where this is beneficial, e.g. infections by pathogenic
 CC microorganisms, e.g. bacteria, viruses and protozoans. p62 can be used
 CC to expand T cell populations for treating infectious diseases or cancer,
 CC e.g. the resulting cells may be transduced to render them resistant to
 CC HIV infection. Inhibitors of p62 can be used to reduce B or T cell
 CC responses and may be used to treat a variety of autoimmune diseases,
 CC e.g. diabetes mellitus, arthritis, multiple sclerosis allergic
 CC reactions, Crohn's diseases etc.
 CC
 SQ Sequence 1977 BP: 469 A; 531 C; 546 G; 431 T; 0 other:
 Query Match 1.4%; Score 26; DB 18; Length 1977;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1788 aatcaaaaaaaaaaaaaaatctc 1813
 Db 1939 aatcaaaaaaaaaaaaaaatctc 1964
 RESULT 8
 AAH82014/C
 ID AAH82014 standard; DNA; 251 BP.
 XX
 AC AAH82014;
 XX
 DT 21-SEP-2001 (first entry)
 XX
 DE Rat differential transcription-associated cDNA SEQ ID 523.
 XX
 DE Differential transcription; human; rat; tumour cell; cytostatic;
 KW Ras modulator; Class II tumour suppressor gene; gene therapy; ss.
 XX
 OS Rattus sp.
 XX
 PN W0200157058-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 31-JAN-2001; 2001WO-EP01003.
 XX
 PR 31-JAN-2000; 2000DE-1004102.
 XX
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX
 PI Rosenthal A, Hinzmann B, Schaefer R, Zuber J, Tcherntsa O;
 PI Grips M, Hellriegel M, Schmitz A, Sers C;
 XX
 DR WPI: 2001-483415/52.
 XX
 PT Nucleic acids differentially expressed between tumor and normal cells,
 PT useful for diagnosis or therapy of tumors and for screening active
 PT agents -
 XX
 PS Claim 6: Page 513; 579pp; German.
 XX
 CC This invention describes a nucleic acid (I) with differential expression
 CC between tumour and normal cells and which has cytostatic activity. (I)
 CC work as modulators of Ras activity by inducing expression of tumour

XX (GENY) GENETICS INST INC.
 PA
 XX
 PI Jacobs K, McCoy JM, LaValle ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Bowman MR;
 XX
 DR WPI: 2000-317938/27.
 XX
 PT Isolated polynucleotides, and encoded proteins, comprising secreted
 PT expressed sequence tags (ESTs), useful for treating various disorders
 PT such as autoimmune, infectious, and central nervous system disorders -
 XX
 PS Claim 1: Page 586; 803pp; English.
 XX
 CC AAA43426 to AAA45925 represent specifically claimed secreted expressed
 CC sequence tags (ESTs), isolated from human, mouse, chicken and rat
 CC tissue sources. The ESTs can have a range of activities depending on
 CC the tissues they were isolated from. The activities include:
 CC chemotactic; proliferative; immunomodulatory; haematopoietic;
 CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
 CC cytoskeletal; antibacterial; antifungal; antiviral; antidiabetic;
 CC antiaesthetic; vulnerary; antiulcer; osteopathic; neuroprotective;
 CC neurotropic; antiparkinsonian; antipsoriatic; cerebroprotective;
 CC anticonvulsant; and antidepressant. The ESTs can be used for gene
 CC therapy and in vaccines. The ESTs are useful as probes for the
 CC identification and isolation of full-length cDNAs and genomic DNA
 CC molecules which correspond to the ESTs. Proteins encoded by the ESTs
 CC are useful in assays for determining biological activity and raising
 CC antibodies. They may be useful for treatment of autoimmune disorders
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
 CC osteoporosis, osteoarthritis, central nervous system disorders
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
 CC disease), tumours, bacterial, fungal or viral infections, depression and
 CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
 CC in the exemplification of the present invention.
 XX
 SO Sequence 579 BP: 109 A; 184 C; 180 G; 106 T; 0 other;

Query Match 8.2%; Score 148; DB 21; Length 579;
 Best Local Similarity 100.0%; Pred. No. 6.9e-44;
 Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1088 ggccttcgagtgaggctggggccatccctggctccctcatgtaagatcttcctctg 1147
 Db 351 GGGCTTCGGGCTGGGCTGGGCCCATCCCTGGCTCCTCATGTCAGACATCTTCCCTCTG 292

QY 1148 catgtcaagagcgtgagcagcagcatctgctctcaccacatgctcatgacctctc 1207
 Db 291 CATGTCAAGGCGCTGGCGACAGGAGCATGCTGCTTCCACCACTGGCTCATGCGCTTTCTC 232

QY 1208 gtgaccaagagatcagcagccatcagg 1235
 Db 231 GTGACCAAGAGGTTGACGACGCTCATGG 204

RESULT 5
 AAF55867 standard; cDNA: 2072 BP.
 ID AAF55867;
 AC AAF55867;
 XX
 XX 17-APR-2001 (first entry)
 DT
 XX
 DE Murine GLUTX1 coding sequence.
 XX
 KM Murine GLUTX1 gene therapy; vaccine; hexose transport modulator;
 KM hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
 KM hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX
 OS Mus sp.

XX
 PN WO200104145-A2.
 XX
 XX 18-JAN-2001.
 PD
 XX
 XX 14-JUL-2000; 2000WO-IB01042.
 XX
 PF 14-JUL-1999; 99US-0143907.
 PR 27-AUG-1999; 99US-0151140.
 PR 23-FEB-2000; 2000US-0184285.
 PR 13-JUL-2000; 2000US-0616132.
 XX
 PA (UYLA-) UNIV LAUSANNE.
 XX
 PI Thorens B, Ibbertson M, Uldry M;
 XX
 DR WPI: 2001-112615/12.
 DR P-PSDB: AAB66934.
 XX
 XX
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischaemia and diabetes -
 XX
 PS Claim 3: Page 73-74; 124pp; English.
 XX
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB6941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is the coding sequence
 CC for murine GLUTX1.
 XX
 SO Sequence 2072 BP: 351 A; 673 C; 584 G; 464 T; 0 other;

Query Match 2.8%; Score 50; DB 22; Length 2072;
 Best Local Similarity 100.0%; Pred. No. 4.5e-09;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 cccatccctgctctcatgcatgagatcttcctctcatgtaagagg 1158
 Db 1191 cccatccctgctctcatgcatgagatcttcctctcatgtaagagg 1240

RESULT 6
 AAF55866 standard; cDNA: 2087 BP.
 ID AAF55866;
 AC AAF55866;
 XX
 XX 17-APR-2001 (first entry)
 DT
 XX
 DE Rat GLUTX1 coding sequence.
 XX
 KM Rat GLUTX1 gene therapy; vaccine; hexose transport modulator;
 KM hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
 KM hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX
 OS Rattus sp.
 XX
 PN WO200104145-A2.
 XX
 XX 18-JAN-2001.
 PD
 XX
 PF 14-JUL-2000; 2000WO-IB01042.
 XX
 PR 14-JUL-1999; 99US-0143907.
 PR 27-AUG-1999; 99US-0151140.
 PR 23-FEB-2000; 2000US-0184285.
 PR 13-JUL-2000; 2000US-0616132.
 XX

CC AAB66933-AAB66941) The GLUTX proteins are related to the facilitative
CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischaemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is the coding sequence
CC for human GLUTX1.

Sequence 2217 BP; 320 A; 750 C; 672 G; 475 T; 0 other;

Query Match	67.68;	Score 1225;	DB 22;	Length 2217;
-------------	--------	-------------	--------	--------------

Best local Similarity 99.58; Pred. No. 0;
Matches 1765; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

[illegible]

Db	1252	aggtagctgttcaacagcgtgtgacgctctccatcaatgagacaagaagcggaaagctgtctcc	1311
OY	903	tgtctctctcaagctgtgtgttcaatgtgtcttcaagcagcagatgtcctctgacgctctaatctcaagc	962
Db	1312	tgtctctctcaagctgtgtgttcaatgtgtcttcaagcagcagatgtcctctgacgctctaatctcaagc	1371
OY	963	tgaaccagaagtgagcccttgacaaatccctcgcacgtgtgacatctcggcgcctgtctctgac	1022
Db	1372	tgaaccagaagtgagcccttgacaaatccctcgcacgtgtgacatctcggcgcctgtctctgac	1431
OY	1023	agccgtcttgagccaagcgtctggagccttgagccttgagcggctggagcaatgtgtccctttca	1082
Db	1432	agccgtcttgagccaagcgtctggagccttgagccttgagcggctggagcaatgtgtccctttca	1491
OY	1083	tgcgcgcgctcttcgagctggagccttgagggcccaatccctctgagctccctcaatgtctcaagatctcc	1142
Db	1492	tgcgcgcgctcttcgagctggagccttgagggcccaatccctctgagctccctcaatgtctcaagatctcc	1551
OY	1143	ctctgcacgtctcaagggcgttggcgcagacagcatctgcctctcaaccaatctgctctcatgtgctct	1202
Db	1552	ctctgcacgtctcaagggcgttggcgcagacagcatctgcctctcaaccaatctgctctcatgtgctct	1611
OY	1203	ttctctgtgaaccaagaagttcctcagcagcctctcatggaagctcctctcaaggccctcatgtgaagctctct	1262
Db	1612	ttctctgtgaaccaagaagttcctcagcagcctctcatggaagctcctctcaaggccctcatgtgaagctctct	1671
OY	1263	ggctctgcctccgactctctgcacatctcaatgtgtctcttcaacttgtctctgtctgagcttgaaa	1322
Db	1672	ggctctgcctccgactctctgcacatctcaatgtgtctcttcaacttgtctctgtctgagcttgaaa	1731
OY	1323	cttaaaagaaagactctctgaaacaatctacaagcccatcttgaggggcgtatgtacagccaatca	1382
Db	1732	cttaaaagaaagactctctgaaacaatctacaagcccatcttgaggggcgtatgtacagccaatca	1791
OY	1383	ctaaaggagatlgagcagaagccttggactccaagccttgagcccaaggcccaaggccctctgctgc	1442
Db	1792	ctaaaggagatlgagcagaagccttggactccaagccttgagcccaaggcccaaggccctctgctgc	1851
OY	1443	cccaagggaagccaagaatcccaagccctcttgagagccttgctcagaggtccctccctctcgt	1502
Db	1852	cccaagggaagccaagaatcccaagccctcttgagagccttgctcagaggtccctccctctcgt	1911
OY	1503	catgtctccctccagcccaatgacccgggctcaggaagctcaactgtccctctgttcaagctcc	1562
Db	1912	catgtctccctccagcccaatgacccgggctcaggaagctcaactgtccctctgttcaagctcc	1971
OY	1563	tgtctctctctctcgaagactctgaagaaacctcttgaaactcttgagaaactctcgtctgaagccttc	1622
Db	1972	tgtctctctctctcgaagactctgaagaaacctcttgaaactcttgagaaactctcgtctgaagccttc	2031
OY	1623	catgagccaagactaaagcagcaggaaggaagctggagcctctcagaatctcttctctctgc	1682
Db	2032	catgagccaagactaaagcagcaggaaggaagctggagcctctcagaatctcttctctctgc	2091
OY	1683	tggaggtgcttcttgaaagtctgggtgtcgttgagcatctcgtcgtcctctcaagcggagctgct	1742
Db	2092	tggaggtgcttcttgaaagtctgggtgtcgttgagcatctcgtcgtcctctcaagcggagctgct	2151
OY	1743	tatcgggaaggaatattgttctgcacaataaagac	1776
Db	2152	tatcgggaaggaatattgttctgcacaataaagac	2185
RESULT	3		
ID	AA012574		
AC	AA012574 standard; cDNA; 1461 BP.		
XX	AA012574:		
XX	25-SEP-2001 (first entry)		
XX	Human protein having hydrophobic domain encoding cDNA clone HP10784.		
XX			

QY 186 cgggggagatgctgagcgagctgagctgagtgagaccgagcgagcaagctgagcctctgc 245
|||||
Db 290 cgggggagatgctgagcgagctgagctgagtgagaccgagcgagcaagctgagcctctgc 349
QY 246 tgtgctccagctgagcctctgagcgagccttgagcgtatcaaccgagcgagcgagcgagtga 305
|||||
Db 350 tgtgctccagctgagcctctgagcgagccttgagcgtatcaaccgagcgagcgagtga 409
QY 306 tgtgctgagggagcgagcctctcaaccgagccttgagcgtgagtgagtgagtgagcgcc 365
|||||
Db 410 tgtgctgagggagcgagcctctcaaccgagccttgagcgtgagtgagtgagtgagcgcc 469
QY 366 cgggtataatctccgaaatccgcttaaccgagagtcggggggttgctcgagctctgtgagc 425
|||||
Db 470 cgggtataatctccgaaatccgcttaaccgagagtcggggggttgctcgagctctgtgagc 529
QY 426 agctaagtgctgctgagcagcctctgagcagcctgagcagcagcagcagcagcagcagc 485
|||||
Db 530 agctaagtgctgctgagcagcctctgagcagcctgagcagcagcagcagcagcagcagc 589
QY 486 gctgagctgagctgagctgagcagcctctgagcagcctctgagcagcctctgagcagc 545
|||||
Db 590 gctgagctgagctgagcagcctctgagcagcctctgagcagcctctgagcagcctctgagc 649
QY 546 tggccgagagaccccgagcctctgagcagcctctgagcagcctctgagcagcctctgagc 602
|||||
Db 650 tggccgagagaccccgagcctctgagcagcctctgagcagcctctgagcagcctctgagc 709
QY 603 tggagctgctgagcagcctctgagcagcagcagcagcagcagcagcagcagcagcagc 662
|||||
Db 710 tggagctgctgagcagcctctgagcagcagcagcagcagcagcagcagcagcagcagc 769
QY 663 gcttcaacctgagcctgctgagcagcagcagcagcagcagcagcagcagcagcagc 722
|||||
Db 770 gcttcaacctgagcctgctgagcagcagcagcagcagcagcagcagcagcagcagc 829
QY 723 ccttgaatgagccttcaagagcagcagcagcagcagcagcagcagcagcagcagcagc 782
|||||
Db 830 ccttgaatgagccttcaagagcagcagcagcagcagcagcagcagcagcagcagcagc 889
QY 783 tcttgaagagagccaagtctcaagagcagcagcagcagcagcagcagcagcagcagc 842
|||||
Db 890 tcttgaagagagccaagtctcaagagcagcagcagcagcagcagcagcagcagcagc 949
QY 843 aggtgagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 902
|||||
Db 950 aggtgagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1009
QY 903 tggctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 962
|||||
Db 1010 tggctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1069
QY 963 tgaaccagagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1022
|||||
Db 1070 tgaaccagagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1129
QY 1023 agcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1082
|||||
Db 1130 agcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1189
QY 1083 tggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1142
|||||
Db 1190 tggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1249
QY 1143 cctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1202
|||||
Db 1250 cctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1309
QY 1203 tctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1262
|||||
Db 1310 tctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1369
QY 1263 ggcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1322

Db 1370 ggcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1429
QY 1323 ctaagggagagagcctctgagcaaatcaacagccatttgaggggagcagagagcagcagc 1382
|||||
Db 1430 ctaagggagagagcctctgagcaaatcaacagccatttgaggggagcagagagcagcagc 1489
QY 1383 ctaagggagagagagcctctgagcaaatcaacagccatttgaggggagcagagagcagcagc 1442
|||||
Db 1490 ctaagggagagagagcctctgagcaaatcaacagccatttgaggggagcagagagcagcagc 1549
QY 1443 cccagggagagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1502
|||||
Db 1550 cccagggagagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1609
QY 1503 catgctcctccagagccatgagcagcagcagcagcagcagcagcagcagcagcagc 1562
|||||
Db 1610 catgctcctccagagccatgagcagcagcagcagcagcagcagcagcagcagcagc 1669
QY 1563 tgcctgctcctgagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1622
|||||
Db 1670 tgcctgctcctgagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1729
QY 1623 catgagcaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1682
|||||
Db 1730 catgagcaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1789
QY 1683 tggaggtgctttg 1696
Db 1790 tggaggtgctttg 1803

RESULT 2

AAFS5865
ID AAF55865 standard; cDNA: 2217 BP.

AC AAF55865;

DT 17-APR-2001 (first entry)

XX Human GLUTX1 coding sequence.

XX Human; GLUTX; gene therapy; vaccine; hexose transport modulator;

KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;

KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.

OS Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

WP1: 2001-112615/12.
DR P-PSDB: AAB66932.
XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
PT e.g. ischemia and diabetes -
XX
PS Claim 3; Page 70-71; 124pp; English.
XX
CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and

C 961	20	1.1	42	22	AA068926	Mycobacterium tube
C 962	20	1.1	48	16	AA086183	Primer S1Nball17000
C 963	20	1.1	48	17	AAT35073	Sindbis-based, tum
C 964	20	1.1	48	17	AAT30807	Sindbis PCR primer
C 965	20	1.1	48	19	AAV42384	Reverse PCR primer
C 966	20	1.1	48	20	AA297074	Reverse PCR primer
C 967	20	1.1	48	21	AA292785	Sindbis PCR primer
C 968	20	1.1	48	21	AA292912	Sindbis basic vect
C 969	20	1.1	49	22	AAH35472	Sindbis basic vect
C 970	20	1.1	50	19	AAV33498	3' PCR primer used
C 971	20	1.1	51	22	AAH89468	Antisense oligonuc
C 972	20	1.1	51	22	AAH89468	Human coding sequ
C 973	20	1.1	52	18	AAH39252	Human coding sequ
C 974	20	1.1	52	20	AAH58494	Primer 4B used S1N
C 975	20	1.1	54	19	AAV15573	Primer 4B used S1N
C 976	20	1.1	55	15	AA063853	Primer for HIV RNA
C 977	20	1.1	55	21	AA063853	DNA from normal ti
C 978	20	1.1	56	14	AA0472259	Human secreted pro
C 979	20	1.1	56	21	AA0472259	Synthetic polydeox
C 980	20	1.1	56	21	AA0472259	Human secreted pro
C 981	20	1.1	56	21	AA0472259	Human secreted pro
C 982	20	1.1	57	18	AA0472259	Human secreted pro
C 983	20	1.1	58	20	AAH58494	Human secreted pro
C 984	20	1.1	58	20	AAH58494	Primer 11,703R use
C 985	20	1.1	60	22	AAH48688	Primer 11,703R use
C 986	20	1.1	60	22	AAH48688	Human G-protein su
C 987	20	1.1	62	21	AAH48688	Human G-protein su
C 988	20	1.1	63	21	AAH48688	Human secreted pro
C 989	20	1.1	63	21	AAH48688	Arabidopsis thalia
C 990	20	1.1	64	15	AAH48688	Human secreted pro
C 991	20	1.1	64	15	AAH48688	trNApPoly primer.
C 992	20	1.1	65	18	AAH48688	trNApPoly primer. #
C 993	20	1.1	65	20	AAH48688	Reverse primer S1N
C 994	20	1.1	65	21	AAH48688	Reverse primer S1N
C 995	20	1.1	68	21	AAH48688	Human secreted pro
C 996	20	1.1	69	18	AAH48688	Human secreted pro
C 997	20	1.1	74	20	AAV02148	3' portion of cDNA
C 998	20	1.1	74	20	AAV02148	Human secreted pro
C 999	20	1.1	77	21	AAH48688	Primer (ga)10XBNT
C 1000	20	1.1	79	13	AA035096	Human secreted pro
C 1000	20	1.1	79	13	AA035096	Microsatellite seq
C 1000	20	1.1	79	21	AA035096	Human secreted pro

ALIGNMENTS

	RESULT	1
AAD09552	ID	AAD09552 standard; cDNA: 2080 BP.
XX	AC	AAD09552:
XX	DT	10-SEP-2001 (first entry)
XX	DE	Human transporter and ion channel-1 (TRICH-1) cDNA.
XX		Human: transporter and ion channel-1; TRICH-1: vaccine: cystic fibrosis; gene therapy: amyotrophic lateral sclerosis; amnesia: muscular dystrophy; hypertension: angina: neurological disorder: asthma: bipolar disorder: dementia: depression: Alzheimer's disease: epilepsy: mood: arrhythmia: Pick's disease: ischemic cerebrovascular disease; AIDS: anxiety: stroke: Huntington's disease; Parkinson's disease: cerebral neoplasm; allergy: demyelinating disease: mental disorder: Schizophrenia: polymyositis; muscle disorder: cardiomyopathy: cataract: myocarditis: Grave's disease: dermatomycosis; diabetes mellitus: immunological disorder: psoriasis: rheumatoid arthritis: Sjogren's syndrome: systemic lupus erythematosus; sickle cell anemia; Wilson's disease; infertility; Cushing's disease; scleroderma: pulmonary artery stenosis: nocturnal Addison's disease; malabsorption syndrome: hypercholesterolaemia; cancer: ss.
OS		Homo sapiens.
XX		
XX	Key	Location/Qualifiers
FH	CDS	46..1479
TT		

FT		/tag= a
TT		/product= "Human TRICH-1 protein"
XX		
PN	WO200146258-A2.	
XX	28-JUN-2001.	
PD		
XX		
PF	22-DEC-2000; 2000WO-US35095.	
XX		
PR	23-DEC-1999; 99US-0172000.	
XX	14-JAN-2000; 2000US-0176083.	
PR	21-JAN-2000; 2000US-0177332.	
PR	28-JAN-2000; 2000US-0178572.	
PR	02-FEB-2000; 2000US-0179758.	
PR	10-FEB-2000; 2000US-0181625.	
XX		
PA	(INCY-) INCYTE GENOMICS INC.	
XX		
P1	Baught MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy R, Lal P;	
P1	Hillman JL, Azimzal Y, Yue H, Nguyen DB, Yao MG, Gandhi AR;	
XI	Tang YT, Khan FA.	
DR	WPI: 2001-418042/4.	
XX	P-PSDB; AAEO4888.	
PT	Novel human transporter and ion channel proteins useful for treating	
PT	and preventing transporter, neurological, muscle and immunological	
PT	disorders -	
XX		
PS	Claim 5; Page 140-141; 160pp; English.	
XX		
CC	The present sequence is transporter and ion channel-1 (TRICH-1) cDNA.	
CC	TRICH is used as vaccine. TRICH is useful for treating a disease or	
CC	condition associated with decreased expression of functional TRICH,	
CC	such as transport disorder including amyotrophic lateral sclerosis,	
CC	cystic fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth	
CC	disease, Duchenne muscular dystrophy, angina and hypertension,	
CC	neurological disorders including Alzheimer's disease, amnesia, bipolar	
CC	disorder, dementia, depression, epilepsy, ischemic cerebrovascular	
CC	disease, stroke, cerebral neoplasms, Pick's disease, Huntington's	
CC	disease and Parkinson's disease, demyelinating diseases, mental disorders	
CC	including mood, anxiety, Schizophrenia and seasonal affective disorder,	
CC	muscle disorder including cardiomyopathy, myocarditis, polymyositis,	
CC	dermatomyositis, arrhythmias and asthma and immunological disorders	
CC	including AIDS, adult respiratory distress syndrome (ARDS), allergies,	
CC	anemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's	
CC	syndrome, systemic lupus erythematosus and other diseases including	
CC	sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary	
CC	artery stenosis, Grave's disease, Cushing's disease, Addison's disease,	
CC	glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers	
CC	psoriasis and viral, bacterial, fungal, helminthic and protozoal	
CC	infections. TRICH DNA is useful in gene therapy and in diagnostic	
CC	purposes.	
XX		
SQ	Sequence 2080 BP; 300 A; 688 G; 632 G; 460 T; 0 other;	
Query Match	77.2%; Score 1400; DB 22; Length 2080;	
Best Local Similarity	99.8%; Pred. No. 0;	
Matches 1690; Conservative	0; Mismatches 1; Indels 3; Gaps 1	
DY	6 ggcggcgcgcgtcttcctccgcccgcgcgtctcgccgtacgcccttgaggccaccatcagcttgcgt 65	
Db	110 ggagcgcgcgcgtcttcctccgcccgcgcgtctcgccgtacgcccttgaggccaccatcagcttgcgt 169	
DY	66 tcgcgctcgcgtctacagcttcgccgcgcattccatcagcttcagcgcgcgcgcgcgcgcgcgc 125	
Db	170 tcgcgctcgcgtctacagcttcgccgcgcattccatcagcttcagcgcgcgcgcgcgcgcgcgc 229	
DY	126 cgcgccttgagcaaacgcgcgcgcgcgccttcggttcggggcgttgttgacctgggtlccgcggg 185	
Db	230 cgcgccttgagcaaacgcgcgcgcgcgccttcggttcggggcgttgttgacctgggtlccgcggg 289	

659	21	1.2	1589	20	AAx84935	Human secreted pro	742	21	1.2	2019	21	AAF21634	Human breast and o
670	21	1.2	1589	21	AAZ43787	Human fetal brain	743	21	1.2	2078	21	AACT7100	Human ORFX ORF2655
671	21	1.2	1593	20	AAx22249	Human secreted pro	744	21	1.2	2079	22	AAD13169	Human transmembran
C 672	21	1.2	1604	22	AAH18451	Human cDNA sequenc	745	21	1.2	2095	20	AAx27331	Human secreted pro
673	21	1.2	1606	22	AA158222	Human polynucleoti	746	21	1.2	2107	21	AAc35391	Human secreted pro
675	21	1.2	1614	21	AAA29169	wheat 4-hydroxypho	747	21	1.2	2157	20	AA290536	Arabidopsis thalia
676	21	1.2	1616	22	AAD02476	cDNA clone cc71se-	748	21	1.2	2161	22	AAH31408	C33 enzyme encodin
677	21	1.2	1626	21	AAZ98035	Human Hsp-binding	749	21	1.2	2196	22	AAF23897	Human secreted pro
678	21	1.2	1626	22	AAD11648	Human secreted pro	750	21	1.2	2208	21	AAc99137	Human secreted pro
679	21	1.2	1642	21	AAc74242	Human secreted pro	751	21	1.2	2208	22	AAc99137	Human pancreatic c
680	21	1.2	1645	21	AAZ58659	Human secreted pro	752	21	1.2	2270	21	AAV79659	Human secreted pro
681	21	1.2	1651	21	AAc74223	Human interleukin-	C 753	21	1.2	2273	22	AAI62754	Eucalyptus grandis
682	21	1.2	1653	22	AAf74435	Human secreted pro	C 754	21	1.2	2284	22	AAH33306	Human cDNA SEQ ID
C 683	21	1.2	1659	22	AAH24956	Human PRO4 nucleot	C 755	21	1.2	2284	22	AAH14838	Human colon cancer
C 684	21	1.2	1662	21	AAf07629	Nucleotide sequenc	756	21	1.2	2292	19	AAV16361	Human cDNA sequenc
685	21	1.2	1692	21	AAc64559	Fusarium venenatum	757	21	1.2	2299	21	AAf16129	cDNA sequence of t
686	21	1.2	1711	20	AAV84190	Human secreted pro	758	21	1.2	2299	22	AAf72785	Human prostate can
687	21	1.2	1720	20	AAx27439	Sodium channel rec	759	21	1.2	2353	19	AAV45981	Human prostate can
688	21	1.2	1722	19	AAV44898	Human secreted pro	760	21	1.2	2379	21	AAf18204	A. thaliana sterol
689	21	1.2	1723	20	AAV84189	Petunia flower bud	761	21	1.2	2385	21	AAc99048	Lung cancer associ
690	21	1.2	1723	21	AAE22404	Sodium channel rec	762	21	1.2	2407	22	AAH18551	Human TANGO 187 cd
691	21	1.2	1723	21	AAc59418	Human secreted pro	763	21	1.2	2418	21	AAH39979	Human cDNA sequenc
692	21	1.2	1731	21	AAAO2664	Human secreted pro	764	21	1.2	2426	21	AAc50445	Human TANGO 187-2
693	21	1.2	1736	21	AAZ61203	Human colon cancer	C 765	21	1.2	2426	21	AAc50445	Mouse TANGO 240 cd
694	21	1.2	1757	22	AAH16392	cDNA encoding a hu	766	21	1.2	2433	22	AAc98326	Human secreted pro
C 695	21	1.2	1758	19	AAV58786	Human cDNA sequenc	767	21	1.2	2439	21	AAc99048	Human pancreatic c
696	21	1.2	1764	22	AAf26569	Human phospholipas	768	21	1.2	2445	20	AAx87343	Human endodymin cd
697	21	1.2	1765	22	AAH14927	DNA encoding human	769	21	1.2	2477	20	AAZ34287	Human PRO1082 nucl
698	21	1.2	1772	19	AAV34189	Human cDNA sequenc	C 770	21	1.2	2481	19	AAx14086	Human PRO1082 (UNO
699	21	1.2	1772	21	AAZ98102	Human secreted pro	771	21	1.2	2485	21	AAf16016	H. pylori GHP0 359
700	21	1.2	1772	22	AAD11715	Human secreted pro	772	21	1.2	2490	21	AAc39975	Human prostate can
701	21	1.2	1772	22	AAH33010	Human colon cancer	773	21	1.2	2494	22	AAD10112	Human TANGO 187-1
702	21	1.2	1776	20	AAZ32200	Rice hexose carrie	774	21	1.2	2523	21	AAc39978	Rat secreted prote
703	21	1.2	1778	22	AAf54816	Nucleotide sequenc	775	21	1.2	2538	21	AAZ43788	Human TANGO 187-1/
704	21	1.2	1782	21	AAf21846	Human breast and o	776	21	1.2	2562	21	AAc39980	Human fetal brain
705	21	1.2	1784	18	AAf93211	Mouse neurotactin	777	21	1.2	2574	21	AAa27982	Human TANGO 187-3
706	21	1.2	1784	21	AAa08499	Mouse neurotactin	778	21	1.2	2587	17	AAf15761	Human homology to
707	21	1.2	1784	21	AAZ58928	Mouse neurotactin	779	21	1.2	2595	21	AAc39976	Gibberellin (GA1)
708	21	1.2	1784	22	AAf59256	Mouse neurotactin	780	21	1.2	2608	21	AAc74241	Human TANGO 187-2/
709	21	1.2	1804	22	AAf75107	Marine neurotactin	781	21	1.2	2608	20	AAx89196	Human secreted pro
710	21	1.2	1824	20	AAV72463	cDNA encoding a hu	C 782	21	1.2	2660	20	AAx89197	Human DRK1 protei
711	21	1.2	1826	18	AAV04313	Human FGF-15 DNA f	783	21	1.2	2694	22	AAD05625	Human DRK1 antise
712	21	1.2	1826	18	AAf86123	Cotton protoporphy	784	21	1.2	2700	21	AAc39977	Human secreted pro
713	21	1.2	1826	22	AAf76579	Cotton protoporphy	785	21	1.2	2700	22	AAH33268	Human TANGO 187-1/
C 714	21	1.2	1837	19	AAV21647	P. falciparum chor	786	21	1.2	2701	21	AAc53390	Human colon cancer
C 715	21	1.2	1837	21	AAc86980	Plasmodium falcipa	787	21	1.2	2702	19	AAV58772	OB-RGPR2 leptin re
716	21	1.2	1840	22	AAH17520	Human cDNA sequenc	788	21	1.2	2732	21	AAc5937	Plant acetolactate
717	21	1.2	1853	17	AAf38769	Flax susceptible r	789	21	1.2	2752	22	AAf59401	Human TANGO 182 cd
718	21	1.2	1853	21	AAf18001	Lung cancer associ	C 790	21	1.2	2754	19	AAV62746	Human host cell pr
719	21	1.2	1866	21	AAf68010	Eucalyptus grandis	791	21	1.2	2762	22	AAD05588	Human secreted pro
720	21	1.2	1867	22	AAf22434	Human breast cance	C 792	21	1.2	2823	17	AAf35233	Human secreted pro
C 721	21	1.2	1869	22	AAH15827	Human cDNA sequenc	793	21	1.2	2901	22	AAH72836	Natural killer cyt
722	21	1.2	1884	22	AAH34858	Human colon cancer	794	21	1.2	2933	21	AAa62607	Human cervical can
C 723	21	1.2	1889	21	AAc59114	Human secreted pro	795	21	1.2	2935	18	AAf69801	Maize RI cDNA. Ze
C 724	21	1.2	1891	20	AAV69433	Human edg-6 cDNA #	796	21	1.2	2960	22	AAf63590	Aureobasidin sensi
C 725	21	1.2	1891	22	AAH18496	Human cDNA sequenc	C 797	21	1.2	3073	12	AAO14635	Human kidney relat
726	21	1.2	1906	19	AAV64625	Human cDNA sequenc	C 798	21	1.2	3073	17	AAf34371	Clone associated w
727	21	1.2	1912	22	AAf97905	Tomato fructokinase	C 799	21	1.2	3073	20	AAZ32246	Plasmiid pATG29 (AT
728	21	1.2	1931	20	AAZ00477	Human secreted pro	C 800	21	1.2	3073	20	AAZ32246	Human glioblastoma

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 16:25:48 ; Search time 282.42 Seconds
(without alignments)
5503.614 Million cell updates/sec

Title: us-09-516-493-6

Perfect score: 1813
Sequence: 1 aacttgcgcgcgcgcgtct.....aaaaaaaaaaaaattcc 1813

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 930621 seqs, 428662619 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 1000 summaries

Database:

_N_Geneseq_1101:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1400	77.2	2080	22	AAD09552 Human transporter
2	1225	67.6	2217	22	AAE5865 Human GLUTX1 codin
3	589	32.5	1461	22	AAE12574 Human protein expi
4	148	8.2	579	21	AAA44914 Human secreted exp
5	50	2.8	2072	22	AAE55867 Murine GLUTX1 codi
6	41	2.3	2087	22	AAE55866 Rat GLUTX1 coding
7	26	1.4	1977	18	AAH89344 Human p62 cDNA #2.
8	25	1.4	251	22	AAH82014 Rat differential t
9	25	1.4	567	21	AAE08921 Fusarium venenatum
10	25	1.4	1008	21	AAE12934 Aspergillus oryzae
11	25	1.4	1149	21	AAE12939 Aspergillus oryzae

12	25	1.4	1222	20	AAH83994 Human EPC1 encodin
13	25	1.4	3413	21	AAA47459 Human TANGO 239 co
14	25	1.4	3413	21	AAA47483 Human TANGO 239 co
15	24	1.3	249	21	AAE21711 Human breast and o
16	24	1.3	399	20	AAV88775 EST clone HN72. H
17	24	1.3	436	21	AAE79716 Human secreted pro
18	24	1.3	731	13	AAO31693 Real restriction f
19	24	1.3	736	20	AAH90805 Rat progression re
20	24	1.3	747	21	AAE12354 Aspergillus oryzae
21	24	1.3	854	22	AAD05508 Human secreted pro
22	24	1.3	857	22	AAD05554 Human secreted pro
23	24	1.3	1125	21	AAE59477 Human secreted pro
24	24	1.3	1222	20	AAH37475 Human secreted pro
25	24	1.3	1293	21	AAE79177 Human lung tumour-
26	24	1.3	1455	18	AAE98882 Human pancreatic c
27	24	1.3	1711	13	AAQ25817 SCC antigen cDNA.
28	24	1.3	1732	19	AAV34305 Human secreted pro
29	24	1.3	2067	21	AAZ50488 Soybean sulphate p
30	24	1.3	2318	22	AAE24151 Human secreted pro
31	24	1.3	2351	18	AAO62284 Neurotrophic facto
32	24	1.3	2351	18	AAT88210 cDNA for chicken n
33	24	1.3	3212	22	AAE84373 Nucleotide sequenc
34	24	1.3	3861	21	AAZ93173 Polynucleotide enc
35	24	1.3	3861	21	AAZ33353 Human secreted pro
36	24	1.3	3967	22	AAE07041 Saccharomyces cere
37	24	1.3	3967	22	AAD03830 Saccharomyces cere
38	24	1.3	10545	21	AAE69132 Human ABC1 gene ex
39	24	1.3	17041	13	AAQ21065 Genomic DNA of hum
40	24	1.3	17041	13	AAQ67057 Human PACAP genom
41	24	1.3	84607	20	AAE90847 Human PACAP genom
42	24	1.3	183999	22	AAE92831 A. thaliana gene a
43	23	1.3	38	20	AAZ23096 Flea saliv proteol
44	23	1.3	420	19	AAE73419 Flea saliv proteol
45	23	1.3	420	19	AAV73420 Flea saliv proteol
46	23	1.3	484	21	AAE10693 Human hypoxia resp
47	23	1.3	484	21	AAE251563 Human hypoxia resp
48	23	1.3	567	22	AAH68735 Human cervical can
49	23	1.3	572	22	AAI41139 Probe #825 used t
50	23	1.3	579	22	AAH73226 Human cervical can
51	23	1.3	580	21	AAE25162 Human hypoxia resp
52	23	1.3	595	22	AAH87993 Peppermint plant o
53	23	1.3	595	22	AAH70163 Human secreted pro
54	23	1.3	599	22	AAH32554 Human secreted pro
55	23	1.3	599	22	AAH32554 Flea saliv proteol
56	23	1.3	606	19	AAV73439 Human inflammatory
57	23	1.3	606	19	AAH92265 Human inflammatory
58	23	1.3	700	22	AAH92265 Human secreted pro
59	23	1.3	728	21	AAE34484 Human secreted pro
60	23	1.3	739	21	AAE35252 Ubiquitome oxidore
61	23	1.3	837	21	AAE08144 Fusarium venenatum
62	23	1.3	922	22	AAH34604 Human colon cancer
63	23	1.3	998	18	AAE79941 Full length chitin
64	23	1.3	1094	22	AAE92125 Human PRO1926 cDNA
65	23	1.3	1221	20	AAE41376 Human normal uteru
66	23	1.3	1297	22	AAE97953 Human secreted pro
67	23	1.3	1346	21	AAE40664 Human serine/threo
68	23	1.3	1361	21	AAE98062 Human colon cancer
69	23	1.3	1588	19	AAV43614 Human secreted pro
70	23	1.3	1718	21	AAE61418 DNA encoding the s
71	23	1.3	1718	21	AAZ50887 Human LDCAM bindin
72	23	1.3	1737	22	AAE50887 Abscisic acid resp
73	23	1.3	1737	22	AAE50887 Abscisic acid resp
74	23	1.3	1741	22	AAE50887 Arabidopsis thalia
75	23	1.3	1741	22	AAE50887 Arabidopsis thalia
76	23	1.3	1766	18	AAE91332 Abscisic acid resp
77	23	1.3	1766	18	AAE91332 Arabidopsis thalia
78	23	1.3	1819	20	AAE01456 A. thaliana Inorga
79	23	1.3	1820	21	AAE61416 Clone np206.8 enco
80	23	1.3	1820	21	AAE50886 DNA encoding the l
81	23	1.3	2026	21	AAE50886 Human LDCAM bindin
82	23	1.3	2027	19	AAE25979 Human pancreatic c
83	23	1.3	2079	21	AAE25979 Human CD33-like pr
84	23	1.3	2150	21	AAE26360 Human secreted pro

/protein_id="CAC28496.1"
/db_xref="GI:12711201"
/translation="MSPEDPQETOPLLRSPGARGRVFLATFAALGPLSPFAL
GYSSPAIPSLRTPALRLGDTAASWFGAVTLGAAGVILGMLDRGRKLSLL
CYVPFVFAVITTAARDVWMLGGRLITGACCVASLVAVYISETIYPRVGLISGC
VOLMVTGILATVAGVLEMRNLAVIGCPPTLMILMKYMETPTPEFLTQHOYDA
MAALRLWSEEGEERPVGAHOGFOLALRRGVKRLPLITGICLWVFOQLSVNAI
MFAANTIFEEAKFKDSSLASVTVGIIQVLETAVALIMDRAGRLLLALGVIMVFSM
SAEGYFKLQSGPSSSHVGLVPISSAEPADVHLGLAWLAVGSMCLFIAGFAWGP
IPALLMSEIFPLIKGVATGCVLTNNMFALVTKERNIMEILRPYGAFMLTAFCI
LSVLEFLTFEPETKGRLEQITAHPEGR"

BASE COUNT 355 a 673 c 580 g 479 t

ORIGIN

Query Match 2.3%; Score 41; DB 6; Length 2087;
Best Local Similarity 100.0%; Pred. No. 5e-10;
Matches 41: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 cccatccctggctcctcatgtcagagatcttcctctgca 1149
|||||
Db 1203 CCCATCCCCGTGCTCCTCATGTCAAGATCTTCCCTCTGCA 1243

RESULT 15
RNO245935 2087 bp mRNA ROD 18-FEB-2000
LOCUS Rattus norvegicus mRNA for glucose transporter (GLUTX1 gene).
DEFINITION AJ245935
ACCESSION AJ245935.1 GI:7018604
VERSION
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 2087)
Ibberson,M., Uldry,M. and Thorens,B.
GLUTX1, a novel mammalian glucose transporter expressed in the
central nervous system and insulin-sensitive tissues
J. Biol. Chem. 275 (7), 4607-4612 (2000)
2 (bases 1 to 2087)
Ibberson,M.R.
Direct Submission
Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
CH-1005, SWITZERLAND
location/Qualifiers
1. 2087
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
30.1466
/gene="GLUTX1"
30.1466
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CA875729.1"
/db_xref="GI:7018605"

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS

BASE COUNT 355 a 673 c 580 g 479 t

ORIGIN

Query Match 2.3%; Score 41; DB 10; Length 2087;
Best Local Similarity 100.0%; Pred. No. 5e-10;
Matches 41: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 cccatccctggctcctcatgtcagagatcttcctctgca 1149
|||||
Db 1203 CCCATCCCCGTGCTCCTCATGTCAAGATCTTCCCTCTGCA 1243

Search completed: February 13, 2002, 20:05:17
Job time: 22934 sec

BASE COUNT 351 a 673 c 584 g 464 t
ORIGIN
PMLMSEIFPLHVGATGTCGTCATGTCAGAGATCTTCCCTGTCATGTCAGAGG 1240
SVLEFLLVPEETGRTLEQVTAHFEGR

Query Match 2.8%; Score 50; DB 6; Length 2072;
Best Local Similarity 100.0%; Pred. No. 9e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1109 cccatccctgctctcatgctcaagatctccctgcgcatgcaagg 1158
|||||
Db 1191 CCCATCCCTGCTCCTCATGTCAGAGATCTTCCCTGTCATGTCAGAGG 1240

RESULT 12
LOCUS MMU245936 2072 bp mRNA ROD 18-FEB-2000
DEFINITION Mus musculus mRNA for glucose transporter (GLUTX1 gene).
ACCESSION AJ245936
VERSION AJ245936.1 GI:7018606
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2072)
AUTHORS Ibberson, M., Uldry, M. and Thorens, B.
TITLE GLUTX1, a novel mammalian glucose transporter expressed in the central nervous system and insulin-sensitive tissues
JOURNAL J. Biol. Chem. 275 (7), 4607-4612 (2000)
MEDLINE 20138191
REFERENCE 2 (bases 1 to 2072)
AUTHORS Ibberson, M.R.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne, CH-1005, SWITZERLAND

FEATURES
SOURCE location/Qualifiers
1..2072
/organism="Mus musculus"
/db_xref="taxon:10090"
21..1454
/gene="GLUTX1"
21..1454
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CA875719.1"
/db_xref="GI:7018607"

/translation="MSPEDDPOETPLLRPPKRPGRVFLASFAALGFLNFGPAL
GYSSPAIPSLRTRAPALRUGDNASWFGAVVTGAAGGILGGLDRLSRKLSLL
CTVPEVFAVITAARVWMLLGRLLTGACGVAIVAPVISEIYPAVAGLIGSC
VOLAVVIGILAVAGWVLEWMLAVGCVPTMLMLCMTPTPFLPLTHOYQEA
MAALRPLMGSEEGWDEPPVGAHOGFOLALRPGIKKPLIGISLWFOOLSGVNAI
MEANRIFEEAKFEDSSLASVTVGIIIOVLPAVALIMDRAGRLLALSGVIVFPM
SAGTVYKTIQSLPSNSHVGIVPIAEPDVQGLAMLAGSCFLTAGVAVMGPI
PMLMSEIFPLHVGATGTCGTCATGTCAGAGATCTTCCCTGTCATGTCAGAGG
SVLEFLLVPEETGRTLEQVTAHFEGR
BASE COUNT 351 a 673 c 584 g 464 t
ORIGIN

Query Match 2.8%; Score 50; DB 10; Length 2072;
Best Local Similarity 100.0%; Pred. No. 9e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1109 cccatccctgctctcatgctcaagatctccctgcgcatgcaagg 1158
|||||
Db 1191 CCCATCCCTGCTCCTCATGTCAGAGATCTTCCCTGTCATGTCAGAGG 1240

RESULT 13

AF321324 AF321324 1012 bp mRNA MAM 02-JUL-2001
LOCUS Bos taurus glucose transporter 8 (Glut8) mRNA, partial cds.
DEFINITION AF321324
ACCESSION AF321324
VERSION AF321324.1 GI:14582715
KEYWORDS
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 1012)
AUTHORS Augustin, R., Navarrete-Santos, A. and Fischer, B.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2000) Anatomy and Cell Biology, Faculty of Medicine, Martin Luther University Halle, Grosse Steinstrasse 52, Halle 061097, Germany

FEATURES
SOURCE location/Qualifiers
1..1012
/organism="Bos taurus"
/db_xref="taxon:9913"
<1..1012
/gene="Glut8"
<1..1005
/gene="Glut8"
/codon_start=1
/product="glucose transporter 8"
/protein_id="AAK69606.1"
/db_xref="GI:14582716"

/translation="ISEIAYPEVGVGLGSCVOLVVGILLAYLAGWLEWMLAVLS
CVAPSEMLLMCFMPEETPRLSOHKHOEAMAMQFLMGVKGWEEPTLGHODFHY
AOLRRPGVYKPEITIGISLMAFOOLSGVAVVAFETFEAKFEDSSLASVVGIVQY
LEFATPALIMDRAGRLLITSGVVMVPSVAPGTYEKLTEGCSNSSHNDLPALYSM
EADITNGIAMLAAGNMCLITAGVAVGMPILPMLMSEIFPLHVGATGTCATGTCAGAGG
MAFLVTKFESSLMEVLRPYGAFLASACIFGVLTACVPEETGRTLEQVTAHFEGR

BASE COUNT 153 a 323 c 302 g 234 t
ORIGIN
Query Match 2.6%; Score 47; DB 4; Length 1012;
Best Local Similarity 100.0%; Pred. No. 3.7e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1118 tggctctcatgctcaagatctccctgcgcatgcaaggcgctggc 1164
|||||
Db 751 TGGCTCCTCATGTCAGAGATCTTCCCTGTCATGTCAGAGGCGCTGGC 797

RESULT 14
LOCUS AX076669 2087 bp DNA PAT 06-FEB-2001
DEFINITION Sequence 3 from Patent WO0104145.
ACCESSION AX076669
VERSION AX076669.1 GI:12711200
KEYWORDS

SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 2087)
AUTHORS Thorens, B., Ibberson, M. and Uldry, M.
TITLE Glutx polypeptide family and nucleic acids encoding same
JOURNAL Patent: WO 0104145-A 3 18-JAN-2001;
University of Lausanne (CH)

FEATURES
SOURCE location/Qualifiers
1..2087
/organism="Rattus sp."
/db_xref="taxon:10118"
30..1466
/note="unnamed protein product"
/codon_start=1

MMU17802 1490 bp mRNA ROD 13-MAY-2000
 LOCUS MMU17802 1490 bp mRNA for glucose transporter 8 (GLUT8 gene).
 DEFINITION Y17802.1 GI:7688219
 ACCESSION
 VERSION
 KEYWORDS glucose transporter 8; GLUT8 gene.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1490)
 AUTHORS Doege, H., Schumann, A., Bahrenberg, G., Brauers, A. and Joost, H. G.
 TITLE GLUT8, a novel member of the sugar transport facilitator family with glucose transport activity
 JOURNAL J. Biol. Chem. 275 (21), 16275-16280 (2000)
 MEDLINE 20283667
 REFERENCE 2 (bases 1 to 1490)
 AUTHORS Joost, H. G.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUL-1998) H. Joost, Institute of Pharmacology and Toxicology, Technical University Aachen, Wendlingweg 2, D-52057 Aachen, FRG
 FEATURES
 source Location/Qualifiers
 1..1490
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /tissue_type="testis"
 26..1459
 /gene="GLUT8"
 26..1459
 /gene="GLUT8"
 /codon_start=1
 /product="glucose transporter 8"
 /protein_id="CA89815.1"
 /db_xref="GI:7688220"
 /translation="MSPEDPQETOPLLRPPPARPPGRRVFLASFAALGPIISGFAL
 GYSSPAIPSLRRTAPPALRLGDNAASFVAVITLGAAGITLGGMLDRGKRKLSLL
 CTPEVYGFVITTAARDVWMLLGGRLTLGACGVAISLAVYISEIAYPAVRGLGSC
 VOLMVYIGILLAYVAGVWLEMRMLAVGCVPTMLLMCMPTPEPFLTHOYORA
 MAALRLFMGSEEGMEPPVGAEHOGFOLALRRRGYKPLIIGISLWVFOOLSGVNI
 MFYANSIFFEAKFKRDSLSAVYGIIOVLTFAVAALIMDRAGRLLALAGSLVWFSM
 SAFGTYFKLTQSLPSNSSHVGIVPIAEPPDVYGLAMLAVGSKCLFIAGFVWGPI
 PWLMSIFPLHVGATGICVLINMFMALVTKEFSSVMELRPYGAFWLTAFCAL
 SVLTFLVPEPKGRLEQVTAHFEGR"
 BASE COUNT 222 a 499 c 438 g 331 t
 ORIGIN
 Query Match 2.88; Score 50; DB 10; Length 1490;
 Best Local Similarity 100.0%; Pred. No. 9.3e-15;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1109 cccatccctgcgtccatcgtcagagatcttcctcgtcgtcagaggg 1158
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1196 CCCATCCCTGCTCCTCATGTCAAGATCTCCCTCGCATGTCAAGGG 1245
 RESULT 10
 AF232061 1843 bp mRNA ROD 23-JUN-2000
 LOCUS AF232061 1843 bp mRNA for glucose transporter GLUT8 mRNA, complete cds.
 DEFINITION AF232061
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1843)
 AUTHORS Carayannopoulos, M. O., Chi, M. M., Cui, Y., Pingsterhaus, J. M.,
 McNight, K. A., Mueckler, M., Devaskar, S. U. and Moley, K. H.
 TITLE GLUT8 is a glucose transporter responsible for insulin-stimulated
 glucose uptake in the blastocyst

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (13), 7313-7318 (2000)
 MEDLINE 20319023
 REFERENCE 2 (bases 1 to 1843)
 AUTHORS Moley, K. H., Carayannopoulos, M. O. and Cui, Y.
 TITLE Direct Submission
 JOURNAL Submitted (07-FEB-2000) OB/Gyn, Washington University, 4566 Scott
 Ave, St. Louis, MO 63110, USA
 FEATURES
 source Location/Qualifiers
 1..1843
 /organism="Mus musculus"
 /strain="129"
 /db_xref="taxon:10090"
 1..48
 49..1482
 /codon_start=1
 /product="glucose transporter GLUT8"
 /protein_id="AA878366.1"
 /db_xref="GI:8671758"
 /translation="MSPEDPQETOPLLRPPPARPPGRRVFLASFAALGPIISGFAL
 GYSSPAIPSLRRTAPPALRLGDNAASFVAVITLGAAGITLGGMLDRGKRKLSLL
 CTPEVYGFVITTAARDVWMLLGGRLTLGACGVAISLAVYISEIAYPAVRGLGSC
 VOLMVYIGILLAYVAGVWLEMRMLAVGCVPTMLLMCMPTPEPFLTHOYORA
 MAALRLFMGSEEGMEPPVGAEHOGFOLALRRRGYKPLIIGISLWVFOOLSGVNI
 MFYANSIFFEAKFKRDSLSAVYGIIOVLTFAVAALIMDRAGRLLALAGSLVWFSM
 SAFGTYFKLTQSLPSNSSHVGIVPIAEPPDVYGLAMLAVGSKCLFIAGFVWGPI
 PWLMSIFPLHVGATGICVLINMFMALVTKEFSSVMELRPYGAFWLTAFCAL
 SVLTFLVPEPKGRLEQVTAHFEGR"
 BASE COUNT 291 a 609 c 529 g 410 t 4 others
 ORIGIN
 Query Match 2.88; Score 50; DB 10; Length 1843;
 Best Local Similarity 100.0%; Pred. No. 9.1e-15;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1109 cccatccctgcgtccatcgtcagagatcttcctcgtcgtcagaggg 1158
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1219 CCCATCCCTGCTCCTCATGTCAAGATCTCCCTCGCATGTCAAGGG 1268
 RESULT 11
 AX076671 2072 bp DNA PAT 06-FEB-2001
 LOCUS AX076671 2072 bp DNA
 DEFINITION Sequence 5 from Patent WO0104145.
 ACCESSION AX076671
 VERSION AX076671.1 GI:12711202
 KEYWORDS
 SOURCE Mus sp.
 ORGANISM Mus sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2072)
 AUTHORS Thorens, B., Ibbersson, M. and Udry, M.
 TITLE Glutx polypeptide family and nucleic acids encoding same
 JOURNAL Patent: WO 0104145-A 5 18-JAN-2001;
 University of Lausanne (CH)
 FEATURES
 source Location/Qualifiers
 1..2072
 /organism="Mus sp."
 /db_xref="taxon:10095"
 21..1454
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC28497.1"
 /db_xref="GI:12711203"
 /translation="MSPEDPQETOPLLRPPPARPPGRRVFLASFAALGPIISGFAL
 GYSSPAIPSLRRTAPPALRLGDNAASFVAVITLGAAGITLGGMLDRGKRKLSLL
 CTPEVYGFVITTAARDVWMLLGGRLTLGACGVAISLAVYISEIAYPAVRGLGSC
 VOLMVYIGILLAYVAGVWLEMRMLAVGCVPTMLLMCMPTPEPFLTHOYORA
 MAALRLFMGSEEGMEPPVGAEHOGFOLALRRRGYKPLIIGISLWVFOOLSGVNI
 MFYANSIFFEAKFKRDSLSAVYGIIOVLTFAVAALIMDRAGRLLALAGSLVWFSM
 SAFGTYFKLTQSLPSNSSHVGIVPIAEPPDVYGLAMLAVGSKCLFIAGFVWGPI

QY 126 cgcgcctgagacgacgcgcgcctctctggtctcgggagctgctgacccctgagtcgcgcg 185
|||||
DB 185 CGCGCTGAGACGACGCCCGCTCTGTTGCGGGGCTGTCTGACACCGCTGGCTGCCGG 244
QY 186 cgggggagatctctgagcgctgctgctgctgctgctgctgctgctgctgctgctgctg 245
|||||
DB 245 CGGGGGAGTCTGGGCGGCTGGCTGGGACCGCGGGGCGCAAGCTGACCTCTTTC 304
QY 246 tctctcctgctcctctctgctgctgctgctgctgctgctgctgctgctgctgctgctg 305
|||||
DB 305 TGTGCTCCGCTCCCTCTGCGGCGGCTTTCGCTCATCACCGCGGCGCAAGAGCTGTGGA 364
QY 306 tctctgctgagggcgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 365
DB 365 TGTGCTGGGGGCGCGCTCTCTCACCGGCGCTGGCTGGGCTGTGCTCTCTCTCTCTCTCT 424
QY 366 cgtctacatctcgaatctcgaatctcgaatctcgaatctcgaatctcgaatctcgaatctc 425
|||||
DB 425 CGGCTACATCTCCGAAATCGCTACCGACAGTCGGGGGCTTGTCTGCTCTCTCTCTCTCT 484
QY 426 agctaagtctcgtctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 485
|||||
DB 485 AGCTAATGCTGCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 544
QY 486 gctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 545
|||||
DB 545 GCTGGCTGGCTGCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 604
QY 546 tgcgcgagaccccgctctctctgctgctgctgctgctgctgctgctgctgctgctgctg 594
|||||
DB 605 TGCCCGAGACCCCGGCTTCTGCTGACTCAGACAGCGCGCAGAGAGG 653

RESULT 6
AX191507 1461 bp DNA PAT 15-AUG-2001
LOCUS AX191507
DEFINITION Sequence 29 from Patent WO0149728.
ACCESSION AX191507
VERSION AX191507.1 GI:15209697
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1461)
AUTHORS Kato, S. and Kimura, T.
TITLE Human proteins having hydrophobic domains and dnas encoding these
JOURNAL Proteins
Patent: WO 0149728-A 29 12-JUL-2001;
Proteome Inc. (JP), SAGAMI CHEMICAL RESEARCH CENTER (JP)
FEATURES
source Location/Qualifiers
1..1461
/organism="Homo sapiens"
/db_xref="taxon:9606"
61..849
/note="unnamed protein product"
/protein_id="CAC51153.1"
/codon_start=1
/db_xref="GI:15209698"
CDS
translation="MTPDEPEETPLPGPGSGAPKGRHVLAFAPALGFLSTGFL
GYSPPAIPSLQRAAPAPRILDDAASWFGAVITGAAGVGLGVLDRARSLSL
CSVPFVAGFAVITAADVWMLIGRLITGLAGVSAIVAPYISIEIVAPVAGLIGSC
VOLNMGILATLAGWLEWRLAVLGCVPISLMLMCPMPREPLRLQHRRQEA
APGLVRGSHGQHRCRLRLQADPGMPWOLLARHGLGCLCTTAC"

BASE COUNT 195 a 501 c 456 g 309 t
ORIGIN

Query Match 32.5%; Score 589; DB 6; Length 1461;
Best Local Similarity 100.0%; Pred. No. 4.5e-299;
Matches 589; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ggcgcgcgcgcgtctctcgcgcgcctcgcgcctgagccactcagctcgcgt 65

DB 125 CGGGCGCGCGCTCTCTCTGCGGCGCTTTCGCGCTGCGCTTGGGCGCTACACTTCGCGCT 184
QY 66 tgcgcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 125
DB 185 TCGCGCTGAGCTACAGCTTCCCGGCTATCCCTAGCTGAGAGCGCGCGCGCGCGCGCG 244
QY 126 cgcgcctgagacgacgcgcgcctcctctggtctcgggagctgctgacccctgagtcgcgcg 185
|||||
DB 245 CGCGCTGAGACGACCGCGCGCTCTCTGCTGGGAGCTGTCTGACACCGCTGGCTGCCGG 304
QY 186 cgggggagatctctgagcgctgctgctgctgctgctgctgctgctgctgctgctgctg 245
DB 305 CGGGGGAGTCTGGGCGGCTGGCTGGGACCGCGCGGCGCAAGCTGACCTCTTTC 364
QY 246 tctctcctgctcctctgctgctgctgctgctgctgctgctgctgctgctgctgctg 305
|||||
DB 365 TGTGCTCCGCTCCCTCTTGGGCGGCTTTCGCTCATCACCGCGGCGCAAGAGCTGTGGA 424
QY 306 tctctgctgagggcgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 365
|||||
DB 425 TGTGCTGGGGGCGCGCTCTCTCACCGGCTGGCTGGGCTGTGCTCTCTCTCTCTCTCT 484
QY 366 cgtctacatctcgaatctcgaatctcgaatctcgaatctcgaatctcgaatctcgaatctc 425
|||||
DB 485 CGGCTACATCTCCGAAATCGCTACCGACAGTCGGGGGCTTGTCTGCTCTCTCTCTCTCT 544
QY 426 agctaagtctcgtctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 485
|||||
DB 545 AGCTAATGCTGCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 604
QY 486 gctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 545
|||||
DB 605 GCTGGCTGGCTGCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 664
QY 546 tgcgcgagaccccgctctctctgctgctgctgctgctgctgctgctgctgctgctgctg 594
|||||
DB 665 TGCCCGAGACCCCGGCTTCTGCTGACTCAGACAGCGCGCAGAGAGG 713

RESULT 7
AL445222 225370 bp DNA PRI 24-APR-2001
LOCUS AL445222
DEFINITION Human DNA sequence from clone RP11-356B19 on chromosome 9, complete
sequence.
ACCESSION AL445222
VERSION AL445222.9 GI:13810082
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 225370)
AUTHORS Laird, G.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk; Clone
requests: clonerequests@sanger.ac.uk
COMMENT On Apr 26, 2001 this sequence version replaced gi:13277497.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given

QY 246 tctgctcgtgacccctctgctgagccgtcttgccatccacccgagccagagctgttga 305
 |||||
 Db 652 tctgctcgtgacccctctgctgagccgtcttgccatccacccgagccagagctgttga 711
 |||||
 QY 306 tctgctcgtgagccgtctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 365
 |||||
 Db 712 tctgctcgtgagccgtctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 771
 |||||
 QY 366 cgtgctcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 425
 |||||
 Db 772 cgtgctcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 831
 |||||
 QY 426 agcctcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 485
 |||||
 Db 832 agcctcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 891
 |||||
 QY 486 gctgctcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 545
 |||||
 Db 892 gctgctcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 951
 |||||
 QY 546 tgcctcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 602
 |||||
 Db 952 tgcctcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1011
 |||||
 QY 603 tgcctcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 662
 |||||
 Db 1012 tgcctcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1071
 |||||
 QY 663 gcttccacgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 722
 |||||
 Db 1072 gcttccacgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1131
 |||||
 QY 723 ccttccacgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 782
 |||||
 Db 1132 ccttccacgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1191
 |||||
 QY 783 tcttccacgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 842
 |||||
 Db 1192 tcttccacgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1251
 |||||
 QY 843 aggtcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 902
 |||||
 Db 1252 aggtcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1311
 |||||
 QY 903 tcttccacgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 962
 |||||
 Db 1312 tcttccacgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1371
 |||||
 QY 963 tgcctcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1022
 |||||
 Db 1372 tgcctcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1431
 |||||
 QY 1023 aggtcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1082
 |||||
 Db 1432 aggtcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1491
 |||||
 QY 1083 tgcctcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1142
 |||||
 Db 1492 tgcctcgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1551
 |||||
 QY 1143 tcttccacgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1202
 |||||
 Db 1552 tcttccacgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1611
 |||||
 QY 1203 tcttccacgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1262
 |||||
 Db 1612 tcttccacgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1671
 |||||
 QY 1263 gcttccacgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1322
 |||||
 Db 1672 gcttccacgtgacccctctgacccgtctgagccgtctgagccgtctgagccgtctgagcc 1731
 |||||
 QY 1323 ctaaggaagactctgtaacaatacagccattttagagggcgatgacagccactca 1382
 |||||

Db 1732 TTAAGGAAGACTCTGGAAACAAATCACAGCCATTGAGGGGGGATGACAGCCACTCA 1791
 |||||
 QY 1383 ctgaggaatgaggaagcaagctgtgactccaagcttgagcccaagcccaagccctgctgac 1442
 |||||
 Db 1792 TTAGGGATGAGAGCAAGCCTGTGACTCCAAGTTGGGGCCCAAGCCCAAGAGCCCTGCTG 1851
 |||||
 QY 1443 cccaaggaagcaagatccagcccttgagcccttgagcccttgagcccttgagcccttgagcc 1502
 |||||
 Db 1852 cccaaggaagcaagatccagcccttgagcccttgagcccttgagcccttgagcccttgagcc 1911
 |||||
 QY 1503 catgctccctcagccatgaccccttgagcccttgagcccttgagcccttgagcccttgagcc 1562
 |||||
 Db 1912 catgctccctcagccatgaccccttgagcccttgagcccttgagcccttgagcccttgagcc 1971
 |||||
 QY 1563 tctgctcgtgaggaactccaagcaacccctgagcccttgagcccttgagcccttgagccct 1622
 |||||
 Db 1972 tctgctcgtgaggaactccaagcaacccctgagcccttgagcccttgagcccttgagccct 2031
 |||||
 QY 1623 catgcaagcaagtaagcaagcagcaggaaggaagtgagcccttagatcttgcctcctgac 1682
 |||||
 Db 2032 catgcaagcaagtaagcaagcagcaggaaggaagtgagcccttagatcttgcctcctgac 2091
 |||||
 QY 1683 tgaagtgctttnaggttgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1742
 |||||
 Db 2092 tgaagtgctttnaggttgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 2151
 |||||
 QY 1743 tctcgaaggaagacttcttccaataaagac 1776
 |||||
 Db 2152 TATCGGAAGAAATTTGTTGCCAAATTAAGAC 2185
 |||||

RESULT 4
 HSA17801
 LOCUS HSA17801 1508 bp mRNA PRI 13-MAY-2000
 DEFINITION Homo sapiens mRNA for glucose transporter 8 (GLUT8 gene).
 ACCESSION Y17801
 VERSION Y17801.1 GI:7688145
 KEYWORDS glucose transporter 8; GLUT8 gene.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 AUTHORS Doege,H., Schürmann,A., Bahrenberg,G., Brauers,A. and Joost,H.G.
 TITLE GLUT8, a novel member of the sugar transport facilitator family
 JOURNAL J. Biol. Chem. 275 (21), 16275-16280 (2000)
 MEDLINE 20283667
 REFERENCE 2 (bases 1 to 1508)
 AUTHORS Joost,H.G.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUL-1998) H. Joost, Institute of Pharmacology and
 Toxicology, Technical University Aachen, Wendlingweg 2, D-52057
 Aachen, FRG

FEATURES
 source
 1..1508
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="testis"
 27..1460
 /gene="GLUT8"
 /gene="GLUT8"
 /codon_start=1
 /product="glucose transporter 8"
 /protein_id="CAB89809.1"
 /db_xref="GI:7688146"
 /translation="MTPEDEPTQPLAGPPGSSAPRGRRVFLAFAALGPIISFGL
 GSSPAISLQRAAPARLDADAASMSGAVTTGAAGAGVGLVLRAGKSLIL
 CSVPVAGFAVITTAODVWMLIGRLILGCVGLVAVPYISIAVPAVGLGSGC
 VOLMVVGLITAVIAGVWLEBRWMLVAGCVPSIMLLMCMPEPRFLIQRROEA
 MAALRFMGSDQWDPPIGADQSFHLLALRPGIYKFFIIGVSLMAIQQLSGVAVM

[illegible]

Db	1628	TGCTGCTGCTCTGTAGGACCTTAGGAAACACCTTGCAGCTTTGGACAGACTGGGGTCAACCTC	1687
Oy	1623	catgcgcaagactaaagcaagcggaaagagagtgctcctaagatccttgcctctgcg	16822
Db	1688	CATGCGCAAGACTTAAGCAGCGGAGAGGAGTGGGCGCTCTAGGATCTTTGTCTTCTGGC	1747
Oy	1683	tggaggtgactttgnaaggttgggtgcctgggcaatcgtgcctcctcctaagcgagctcct	1742
Db	1748	TGGAGGTGCTCTTTGGAGGTGGGCTGCGGCATTCGATTCGATTCACACGCGCTGGCT	1807
Oy	1743	catcggaagaaatctgttttgcgaataaagac	1776
Db	1808	TATCGGAGAGGAATTTGTTTGGCAATTAAGC	1841

RESULT	3
AX076667	
LOCUS	AX076667 2217 bp DNA PAT 06-FEB-2001
DEFINITION	Sequence 1 from Patent WO0104145.
ACCESSION	AX076667
VERSION	AX076667.1 GI:12711198
KEYWORDS	.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 2217)
AUTHORS	Thorens,B., Ibberson,M. and Uldry,M.
TITLE	Gltx polypeptide family and nucleic acids encoding same
JOURNAL	Patent: WO 0104145-A 1 18-JAN-2001;
FEATURES	University of Lausanne (CH) location/Qualifiers

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
348..1781
/feature_id="C6C28495.1"
/feature_start=1
/feature_end=1781
/feature_type="CDS"
/feature_qualifiers="
  /note="unnamed protein product"
  /codon_start=1
  /protein_id="C6C28495.1"
  /db_xref="GI:127111199"
  /translation="MPEDPEDEPTEOPLIGPEGGASPRGRVFLAFAAALGPTLSEFEGAL
GYSSPAISLQRAAPAPARLDDAAASFVAVTTEGAAGVGILGMLVDRGRKTLLE
CSYEPFAPVATTAADQVMMLIGRLITGLACGVAISLVAAYVTSIETAPVGRIGSC
VOLMNVVVGILLIATVAGVLEMRRLAVTGCPSPSMLLMLEFMEPTPEILLTHORRODA
MAAIRFLMGSECCHEDEPPICAEOSFHALLRQGIKYKPTITGVSMAFQDSLVNAYM
FTATITTEAKFKDSSLASVAVGTVQLFTVAALLMDRAGRRLILVLSVAVMFTS
AFGAYFKLTGGPGDNSSHVAISAPVSAQPDASVAGLWLAGNMCLFTACFAGWGTI
PMLMSEIFPLHVGAVATGICVLTNMLMAFLVTKFESSLMEVLRPGCAFWLASAPCFE
SVLTSTLCVPEIKGKLTLEQITTAHEGR"
BASE COUNT      320  a      750  c      672  g      475  t
ORIGIN

```

Query Match	67.6%;	Score 1225;	DB 6;	Length 2217;
Best Local Similarity	99.5%;	Pred. No. 0;		
Matches 1765;	Conservative 0;	Mismatches 6;	Indels 3;	Gaps 1.
QY 6	gagagcgcgcgcgtcttcctcgcgcgcgtctgcgcgccttgagcccaactcagcttcgact	65		
Db 412	ggggccggccgctcttccctggccgcttgcgcgccttcgcgccttcgacctgagggccacacacattcggt	471		
QY 66	tcgcgctcgtgcgtacagctctccgcgcacatccctagcctcagcgcgcgcgcgcgcgcgc	125		
Db 472	tgcgccttcgggcacagctcccccggccatccccaagccttcagcgcgcgcgcgcgcgcgc	531		
QY 126	cgcgccttgagcaacgcgcgcgcctcctcgtctcggggcctcgtcgtgaacctcgggtgcgcgg	185		
Db 532	cgcgccttgagcaacgcgcgcgcctcctcctctgcttgggggctctccttgacacctggggtggcgcgg	591		
QY 186	cggggggagctgcctgagcgagcttcgctcgtctgagaccgcgcgcgcgcgaagtcaacctctgc	245		
Db 592	cggggggagctgcctgagccgcttcgctcgtctgagaccgcgcgcgcgcgaagtcaacctcttcgc	651		

```

Oy 843 aggtgctgtcacagactgtgctgctctcatcatgacagagagcgagagctgtctcc 902
    |||
Db 950 AGTGTGCTTTACAGCTGTGGGGCTCTCATCATGACAGAGAGGGCGAGGCTGCTCC 1009
Oy 903 tggctgtgtcaagtgtgtgtcattgtgttcaagcaagatgctgtgagcctacttaagc 962
    |||
Db 1010 TGGTCTTGTGACAGTGTGTGATGTGTGTGACAGCAGATGCTTGGGCTTACTTCAAG 1069
Oy 963 tgaacagaggtgtgacctgtgacctgtgacctgtgacctgtgacctgtgacctgtgacct 1022
    |||
Db 1070 TACACGAGGAGTGTGCTTGGCACTCTCCACAGTGTGCTTGGGCTTGGGCTTGGGCTTGG 1129
Oy 1023 agcctgtgtgtacagcagctgtgagagctgtgacctgtgacctgtgacctgtgacctgt 1082
    |||
Db 1130 AACCTGTGTATGCCAGCGCTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1189
Oy 1083 tgcgcgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1142
    |||
Db 1190 TGGCGGCTTTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1249
Oy 1143 ctctgcatgtcaagagcgtgtgagcagagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1202
    |||
Db 1250 CTCTGCTATTCAGAGGCGTGTGGCAGACAGCATGTGCTTCTTCACTTCACTTCACTTCA 1309
Oy 1203 tctctgtgtcaagagagctgtgagcagcctcaatgtgtgtgtgtgtgtgtgtgtgtgt 1262
    |||
Db 1310 TTTCTGTGTACCAAGAGGTTTACAGAGCCTCATGTGAGTGTCTTCAAGGCTTATGAGGCT 1369
Oy 1263 ggcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1322
    |||
Db 1370 GGCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1429
Oy 1323 ctaaaggaagagactgtgagcaaatcacagacccatttgaagggagatgacagcactca 1382
    |||
Db 1430 CTAAGGAAGAACTGTGGAACAATCACAGCCATTGTGAGGGGCGATGACAGCCACTCA 1489
Oy 1383 ctaagggatgtgagcaagcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1442
    |||
Db 1490 CTAGGGGATGTGAGCAAGCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1549
Oy 1443 ccgaaggagagcagaaatcacagccctgtgagcctgtgtgtgtgtgtgtgtgtgtgtgt 1502
    |||
Db 1550 CCCAGGGAGCCAGAAATCCAGGCCCTTGGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1609
Oy 1503 catgtctccctcagcccaatgacccgaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1562
    |||
Db 1610 CATGTCTCCCTCAGGCCATGACCCGGGGCTAGAGAGGCTCACTGCTTGTGTGTGTGTGT 1669
Oy 1563 tgcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1622
    |||
Db 1670 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1729
Oy 1623 catgtcgaagactaaagcagcgaagaagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1682
    |||
Db 1730 CATGCCCAAGACTAAAGCAGCGGAGAGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1789
Oy 1683 tgaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1696
    |||
Db 1790 TGGAGGTGCTTTTG 1803

```

```

AUTHORS      Ibberson,M., Uldry,M. and Thorens,B.
TITLE        GLUTX1, a novel mammalian glucose transporter expressed in the
              central nervous system and insulin-sensitive tissues
JOURNAL      J. Biol. Chem. 275 (7), 4607-4612 (2000)
MEDLINE      20138191
REFERENCE    2 (bases 1 to 1873)
AUTHORS      Ibberson,M.R.
TITLE        Direct Submission
JOURNAL      Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
              and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
              CH-1005, SWITZERLAND

FEATURES
  source      1..1873
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              4..1437
              /gene="GLUTX1"
              4..1437
              /gene="GLUTX1"
              /product="glucose transporter"
              /protein_id="CA875702.1"
              /db_xref="GI:7018306"
              /translation="MPEDEPETHOPLGPPGSGAPRGRRVFLAARALGPSIFRAL
              /GYSPPAIPSLORAAPAPRLDDAASWFAVVTLLAAGVLAGWLVDPARKLSLL
              GSVPEVAGFAVITAADVMMLGLRLTGLAGVSLVAVPVISLTAFAVGLGSLSC
              VOLMVVAGLLAVLAGVLEMLLAVLGVPSLMLLCKFMPETRFLLTQHRROEA
              MAALRFPMSEOGMEDPPIGASGFLALALRPGYIKPTIGVSLMAFPOOLSGVAVM
              FYAETLFEPAKFKDSSLASVVGIVDTFTAAALIMDAGRLLILVSGVMVESTS
              AGATYFKLTGSGPNSHVAISAPSAQVVDVASGLAMVLAVGNMCLFIPAVGNKPI
              PWLMSLIRPLHKGVAITGCVLTMMLMFLVTKERSLMEVLRPIGAWLWASFCIE
              SVLFLCPEIKGRTLRQITLHPEGR"
BASE COUNT   279 a      621 c      573 g      400 t
ORIGIN
Query Match      67.6%; Score 1225; DB 9; Length 1873;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1765; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

```

```

Oy 6 gggagcggcggtcttcctgtgacctgtgacctgtgacctgtgacctgtgacctgtgacctgt 65
    |||
Db 68 GCGGCGCGCGGCTTCTCTGCGCGCTTGTGCGCTGCGCTGCGCGCGCTTGTGCGCTTGTGCG 127
Oy 66 tgcgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 125
    |||
Db 128 TCGCGGCTCGGCTTACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 187
Oy 126 cgcgcctgtgagcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 185
    |||
Db 188 CGCGCTGTGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 247
Oy 186 cggagggaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 245
    |||
Db 248 CGGGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 307
Oy 246 tgtgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 305
    |||
Db 308 TGTGTCTCCGTCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 367
Oy 306 tgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 365
    |||
Db 368 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 427
Oy 366 cggctacatctcgaatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 425
    |||
Db 428 CGGTCTACATCTCGGAATGCTTACCCAGCAGTCGAGGGGTGTGTGTGTGTGTGTGTGTGT 487
Oy 426 agctaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 485
    |||
Db 488 AGCTAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 547
Oy 486 gctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 545

```

```

RESULT 2
LOCUS      HSA245937 1873 bp mRNA PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA for glucose transporter (GLUTX1 gene).
ACCESSION AJ245937
VERSION AJ245937.1 GI:7018305
KEYWORDS   glucose transporter; GLUTX1 gene.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1873)

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 13:43:03 ; Search time 4552.9 Seconds
(without alignments)
6569.304 Million cell updates/sec

Title: US-09-516-493-6
Sequence: 1 aacttgcgcgcgcgcgcgtct.....aaaaaaaaaaaaaatctc 1813

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size : 0
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_hgt_hum:*
31: em_hgt_inv:*
32: em_hgt_rod:*
33: em_hgt_hum:*
34: em_hgt_inv:*
35: em_hgt_rod:*
36: em_hgt_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1400	77.2	2080	6 AX179740	AX179740 Sequence
2	1225	67.6	1873	6 HSA245937	AJ245937 Homo sapi
3	1225	67.6	2217	6 AX076667	AX076667 Sequence
4	1022	56.4	1508	9 HSA17801	Y17801 Homo sapien
5	589	32.5	789	6 AX191497	AX191497 Sequence
6	589	32.5	1461	6 AX191507	AX191507 Sequence
7	463	25.5	225370	6 AL445222	AL445222 Human DNA
8	79	4.4	120	11 G20347	G20347 human STS A
9	50	2.8	1490	10 MMU17802	Y17802 Mus musculu
10	50	2.8	1843	10 AF232061	AF232061 Mus muscu
11	50	2.8	2072	6 AX076671	AX076671 Sequence
12	50	2.8	2072	6 MMU245936	AJ245936 Mus muscu
13	47	2.6	1012	4 AF321324	AF321324 Bos tauru
14	41	2.3	2087	6 AX076669	AX076669 Sequence
15	41	2.3	2087	10 RNO245935	AJ245935 Rattus no
16	41	2.3	2189	10 AB033418	AB033418 Rattus no
17	29	1.6	471	3 AF026265	AF026265 Lucilia c
18	28	1.5	180337	2 AC091546	AC091546 Mus muscu
19	28	1.5	207421	2 AC021643	AC021643 Mus muscu
20	27	1.5	1262	2 BC002443	BC002443 Homo sapi
21	27	1.5	144046	2 AC068334	AC068334 Homo sapi
22	27	1.5	156802	2 AC069527	AC069527 Homo sapi
23	27	1.5	166016	2 AL591378	AL591378 Human DNA
24	27	1.5	175759	10 AF220294	AF220294 Mus muscu
25	26	1.4	1532	9 AF131780	AF131780 Homo sapi
26	26	1.4	1751	3 AF348466	AF348466 Dictyoste
27	26	1.4	1977	6 AR077145	AR077145 Sequence
28	26	1.4	131007	9 HS998C11	AL055106 Human DNA
29	26	1.4	143405	2 AC062023	AC062023 Homo sapi
30	26	1.4	144794	2 AC087632	AC087632 Homo sapi
31	26	1.4	147191	9 AC022234	AC022234 Homo sapi
32	26	1.4	179016	2 CNS05TE1	AL359398 Human chr
33	26	1.4	179016	2 AC087309	AC087309 Homo sapi
34	26	1.4	179237	2 AC026871	AC026871 Homo sapi
35	26	1.4	182648	2 AC076966	AC076966 Homo sapi
36	26	1.4	185097	2 AC023148	AC023148 Homo sapi
37	26	1.4	186534	2 AC091221	AC091221 Homo sapi
38	26	1.4	193999	2 AC048361	AC048361 Mus muscu
39	26	1.4	197514	9 AC018841	AC018841 Homo sapi
40	25	1.4	520	8 PSL127	X70702 P. sativum m
41	25	1.4	534	5 AY027434	X64771 D. discoideu
42	25	1.4	551	3 DDCOXV	AF169407 Mus muscu
43	25	1.4	678	9 BC003602	BC003602 Homo sapi
44	25	1.4	1093	8 AF118382	AF118382 Brassica
45	25	1.4	1207	4 BTU77067	U77067 Bos taurus
46	25	1.4	1365	9 HSM800894	AL110235 Homo sapi
47	25	1.4	1697	9 BC011535	BC011535 Homo sapi
48	25	1.4	1962	10 AF378751	AF378751 Mus muscu
49	25	1.4	1975	10 AF169407	AF169407 Mus muscu
50	25	1.4	2098	3 DDSP60	X52105 Dictyosteal
51	25	1.4	2166	10 AF169408	AF169408 Mus muscu
52	25	1.4	8439	10 AF188624	AF188624 Mus muscu
53	25	1.4	29634	8 YSCL9122	U22383 Saccharomyc
54	25	1.4	78260	2 AC015439	AC015439 Drosophil
55	25	1.4	96565	8 ATT20L15	AL162351 Arabidops
56	25	1.4	110000	2 AC013623	Continuation (3 of
57	25	1.4	110000	2 AC021632	AC021632 Mus muscu
58	25	1.4	138091	2 AC073568	AC073568 Homo sapi
59	25	1.4	157986	2 AC025860	AC025860 Homo sapi
60	25	1.4	159693	2 AC016727	AC016727 Homo sapi
61	25	1.4	167294	2 AC013814	AC013814 Homo sapi
62	25	1.4	172450	2 AC036145	AC036145 Mus muscu
63	25	1.4	179481	2 AC034100	AC034100 Mus muscu
64	25	1.4	183776	2 AL591469	AL591469 Mus muscu
65	25	1.4	190498	2 AC079871	AC079871 Mus muscu
66	25	1.4	200146	2 AC073774	AC073774 Mus muscu
67	25	1.4	202083	2 AC023833	AC023833 Mus muscu
68	25	1.4	234112	2 PEMA1AP2	AL035475 Plasmodiu
69	25	1.4	263776	2 AC087335	AC087335 Mus muscu
70	25	1.4	327927	3 AE003439	AE003439 Drosophil

[illegible]

```

seq_name: /gnt2_6/prodata/2/lna/5A.COMB.seq:US-08-035-928-1
seq_documentation_block:
?
? Sequence 1, Application US/08035928
? Patent No. 5538844
?
? GENERAL INFORMATION:
? APPLICANT: Duyao, Mabel P.
? APPLICANT: MacDonald, Marcy E.
? APPLICANT: Gusella, James F.
? TITLE OF INVENTION: A No. 5538844el Transport Protein Gene from
? TITLE OF INVENTION: the Huntington's Disease Region
? NUMBER OF SEQUENCES: 21
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Sterne, Kessler, Goldstein & Fox
? STREET: 1225 Connecticut Avenue N.W.
? CITY: Washington
? STATE: D.C.
? COUNTRY: U.S.A.
? ZIP: 20036
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/035,928
? FILING DATE: 19930323
? CLASSIFICATION: 435
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 466-0800
? TELEFAX: (202) 833-8716
?
? TELEX:
?
? INFORMATION FOR SEQ ID NO: 1:
?
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1788 base pairs
? TYPE: NUCLEIC ACID
? STRANDEDNESS: both
? TOPOLOGY: linear

```

[illegible]


```

2721 .....GGGGGGTGGTCAACGGGCTCAACCTACTACTGGGCTGCTCTCA. 2677
180 rogluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaIle 196
2676 .....ATGCAGAG..... 2668
197 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspPro...Pr 212
2667 .....TCGCATAGGGAGAGCGCTCGACCCATGCGC 2639
212 oileGlyAla.....GluGlnSerPheHisLeuAlaLeuLeuArgG 226
2638 CTGAGAGCCTTCAACCCAGTACGCTCTCCGCTGGCGG..... 2599
226 InProGlyIleThrLysProPheIleIleGlyValSerLeuMetAla... 241
2598 .....CGGGGCACTATCGTCCGCGCACTTATGACTGC 2563
242 .....PheGlnGlnLeuSerGly...ValAsnAlaValMetPheTrpAl 255
2562 TTCTTTATCATGCAACCTGCTAGAGAGGTGCGGCGAGCGCTCTGG.... 2518
255 agluThrIlePheGluGlnAlaLysPheLysAspSer..... 268
2517 ....GTCATTTTCGGGAGAGACCGCTTCGCTGGAGCGGAGCATGATCG 2472
269 ..LeuAlaSerValValGlyValIleGlnValLeuPheThrAla... 283
2471 GCCTGTGCTTGGGTATTCGAAATCTTGCACGCGCTCGCTCAACGCTTC 2422
284 ValAlaAlaLeuIleMetAspArgAlaGlyArgAlaLeuLeuVal... 299
2421 GTCACGTGTCGCCGCCACCAACGTTTCGGCGGAGAGCAAGGCATATATGC 2372
300 .....LeuSerGlyValValMetValPheSerThr... 309
2371 CGGCATGGCGGCGGACGCGCTGGCTACGCTCTGCTGGGCTTCGCGACGC 2322
310 .....SerAlaPheGlyAlaTrpPheLysLeuThrGlnGlyPro 323
2331 GAGGTGATGAGCTTCCCATTAATGATTTCTTCTGCTTCGCGGCGATC 2272
324 GlysSerSerHisValAlaIleSerAlaProValSerAlaGlnProVa 340
2271 GGGATGGCC.....GCGTTGCAGGCGCATGCTGCCAGGCAAGTAGA 2231
340 LAspAlaSerValGly.....LeuAlaTrpLeuAla... 350
2230 TGACGACCATCAGGAGACGCTTCAAGATCGCTCGGCTCTTACCAGCC 2181
351 .....ValGlySerMetCysLeu.....PheIleAlaGlyPhe 361
2180 TAACCTCATCACTGACCGCTGATCGTCACGCGCATTTATGCGGCTCG 2131
362 AlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhe 377
2130 GCGAGCACATGGAAGCGGTGGCATGATTTAGGCGCGGCTTATC 2083
seq_name: /cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-447-430A-40
seq_documentation_block:
: Sequence 40, Application US/08447430A
: Patent No. 5916558
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Recombinant polypeptides and peptides,
: nucleic acids coding for the same and use of these
: TITLE OF INVENTION: polypeptides and peptides in the diagnostic of
: tuberculosis.
: NUMBER OF SEQUENCES: 43
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk

```

```

: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/447,430A
: FILING DATE:
: CLASSIFICATION: 424
: INFORMATION FOR SEQ ID NO: 40:
: SEQUENCE CHARACTERISTICS
: LENGTH: 3423 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: circular
: MOLECULE TYPE: plasmid vector
: HYPOTHEICAL: NO
: US-08-447-430A-40

alignment_scores:
: Quality: 173.00 Length: 416
: Ratio: 0.769 Gaps: 24
: Percent Similarity: 54.087 Percent Identity: 26.202

alignment_block:
US-09-516-493-7 x US-08-447-430A-40/rev ..
Align seg 1/1 to reverse of: US-08-447-430A-40 from: 1 to: 3423

7 AlaPheAlaAlaAlaLeuGlyProLeuSerPhe...GlyPheAlaLeuG 22
||||| : : : ||||| : : : : : : : : : : : : : : : : : : : :
3319 GCGCTCATGCTCACTCTCGGACCGGACCGTACCTGATGCTGTAGCATAGG 3270
22 YTrSerSerProAlaIleProSerLeuGlnArgAlaAlaProAlaP 39
1 : : : : : : : : : : : : : : : : : : : : : : : : : :
3269 CTGTGTTATGCCGTACTGCGGGCGCTCTTCCGGATATGTCATATCC. 3221
39 roArgLeuAspAspAlaAlaIleSerTrpPheGlyAlaValThrLeu 55
||||| : : : ||||| : : : : : : : : : : : : : : : : : : : :
3220 .....GACAGCATGCGCAGTCACTATGCGCTGCTGTAAGCGCTA 3182
56 GlyAla.....AlaAlaGlyValLeuGlyValThrLeuVa 68
||||| : : : : : : : : : : : : : : : : : : : :
3181 TATGCTTATGATGCAATTTCTATGCGACCGCTTCTGAGACA...CTGTC 3135
68 LAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerValProPheV 85
||||| : : : ||||| : : : : : : : : : : : : : : : : : : : :
3134 CGACGCTTGGCGCGCGCCAGTCTGCTGCTGCTGCTGCTGCTGAGACA 3085
85 AlaIleGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeu 101
: : : : : : : : : : : : : : : : : : : : : : : : : :
3084 CTATCGACTACGCGATATGCGGACACACACCGCTCGTGTGATCTCTAC 3035
102 GLYGLYArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAl 118
: : : : : : : : : : : : : : : : : : : : : : : : : :
3034 GCGGAGCATGCTGGCGCGCATCAC...GGGCGCACAGGTGCGGTTCG 2988
118 aproValTrpIleSerGluIleAlaTrpProAlaValArg..... 131
1 : : : : : : : : : : : : : : : : : : : : : : : : : :
2987 TGGCGCTATATGCGCGACATCAGCATGGGGAAGATCGGCGCTCGGCAC 2938
2937 TCGGGCTCATGAGCGCTTGTTCGGGTGGGTGAGTGGCGAGGCGCCGTG 2888
148 LeuAlaTrpLeuAlaGlyTrpVal...LeuGluTrpArgTrpLeuAla 163
: : : : : : : : : : : : : : : : : : : : : : : : : :
2887 GCCGGGAGCTGTTGGCGCATCTCTTGCATGACACCATTCCTTGCG.. 2840
163 lleuGlyCysValProProSerLeuMetLeuLeuMetCysPheMetP 180
: : : : : : : : : : : : : : : : : : : : : : : : : :
2839 ....GGGCGGTCTCAACGCGCTCAACCTACTACTGAGGTGCTCTCA. 2795
180 rogluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaIle 196

```

```
2794 ..... ATGAGAG..... 2786
197 AlaleuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspPro...Pr 212
2785 ..... TCGCATAGGAGAGCGTCGACCCATGCG 2757
212 oLecGlyAla.....GluGlnSerPheHisLeuAlaLeuLeuArgG 226
2756 CTGAGAGCCCTTCAACCCAGTCAGCCTTCGCGTGCGG..... 2717
226 InProGlyIleTyrIleProPheIleIleGlyValSerLeuMetAla... 241
2716 ..... CGGGGATGACTATCGCCGCACTTATGACTGTC 2681
242 ..... PheGlnGlnLeuSerGly...ValAsnAlaValMetPheTyrAl 255
2680 TTCCTTATCATGCAACTGCTAGACAGGTCCGCGACGCTCTGG..... 2636
255 AGluThrIlePheGluGluAlaLysPheLysAspSerSer..... 268
2635 ..... CTCATTTCGGCGGAGACCGCTTCGCTGAGCCGACGATGATCG 2590
269 ..LeuAlaSerValValValGlyValIleGlnValLeuPheThrAla... 283
2389 GCGGTGCGCTTCGCGTATTCGGAATCTTGACGCCCTCCGCAACGCTTCG 2540
284 ValAlaAlaLeuIleMetAspArgAlaGlyArgAlaLeuLeuVal... 295
2539 GTCACTGCTCCCGCCACCAACGCTTTCGGCGAGAACGACGCCATTATCGC 2490
300 ..... LeuSerGlyValValMetValPheSerThr... 309
2489 CGGCATGCGCGCGACGCGCTGGCTGCTGCTGCGGTGCGGAGCC 2440
310 ..... SerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyPro 323
2439 GAGCGTCGATGCGCTCCCATTTATGATTCCTTCGCTCCGCGCGCATC 2390
324 GlysAsnSerSerHisValAlaIleSerAlaProValSerAlaGlnProVa 340
2389 GGGATGCC.....GCGTCGAGCGCATGCTCTCCAGCAGCTGACA 2349
340 LAspAlaSerValGly.....LeuAlaTrpLeuAla... 350
2348 TGACGACCATCAGGACAGCTTCAAGATCGCTCCGCGCTTACAGCC 2299
351 ..... ValGlySerMetCysLeu.....PheIleAlaGlyPhe 361
2298 TAACTTGATCAGTGGACCGCTGATCGTCACGCGCATTTATGCCGCTCG 2249
362 AlaValAlaGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhe 377
2248 GCGAGCAGCATGGAACGGTGTGCGATGTGTGAGCGCGCCCTTATAC 2201
seq_name: /cgn2.6/prodata/2/ina/5b_comb.seq:US-08-447-430A-41
seq_documentation_block:
; Sequence 41, Application US/08447430A
; Patent No. 5916558
; GENERAL INFORMATION:
; APPLICATION:
; TITLE OF INVENTION: Recombinant polypeptides and peptides,
; TITLE OF INVENTION: nucleic acids coding for the same and use of these
; TITLE OF INVENTION: polypeptides and peptides in the diagnosis of
; NUMBER OF SEQUENCES: 43
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
```

```
APPLICATION NUMBER: US/08/447,430A
; FILING DATE:
; CLASSIFICATION: 424
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: plasmid vector
; HYPOTHETICAL: NO
; US-08-447-430A-41

alignment_scores:
  Quality: 173.00      Length: 416
  Ratio: 0.769        Gaps: 24
  Percent Similarity: 54.087      Percent Identity: 26.202

alignment_block:
US-09-516-493-7 x US-08-447-430A-41/rev ..
Align seq 1/1 to reverse of: US-08-447-430A-41 from: 1 to: 3474

7 AlaphAlaAlaAlaLeuGlyProLeuSerPhe...GlyPheAlaLeuG1 22
||||| : : : ||||| : : : : : : : : : : : : : : : : :
3372 GCGCTCATGCTATCCGCGACCGCTACCGCTGAGTGTAGGATAGG 3323
22 yYrSerSerProAlaIleProSerLeuGlnArgAlaIleProProAlaP 39
3322 CTGGTATGACCGGTACTGCGGGGCTCTGGCGGATATCGTCAATTC. 3274
39 roArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValAlaThrLeu 55
||||| : : : ||||| : : : : : : : : : : : : : : : : :
3273 .....GACACATCGCCAGTCACATGACGCGTGTAGGCGCTA 3235
56 GlyAla.....AlaAlaGlyGlyValLeuGlyGlyTrpLeuVa 68
||||| : : : ||||| : : : : : : : : : : : : : : : : :
3234 TATGCGTTGATGCATTTCTATGCGACCGCTTCGGAACA...CTGTC 3188
68 LAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerValProPheV 85
||||| : : : ||||| : : : : : : : : : : : : : : : : :
3187 CGACGCGTTTGGCGCGCGCCAGCTCCTGCTGCTGCTGCTGAGACCA 3138
85 AlaLaglyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeu 101
: : : : : : : : : : : : : : : : : : : : : : : : : : :
3137 CTATGCACTACCGCATGATGAGGACACACACCGCTCCTGATGCTCTAC 3088
102 GlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAl 118
: : : : : : : : : : : : : : : : : : : : : : : : : : :
3087 GCGGAGCGCATCGTGGCGCGCATCAC...GCGGCACAGCTCGGTTGC 3041
118 aproValTyrIleSerGluIleAlaTyrProAlaValArg..... 131
| : : : : : : : : : : : : : : : : : : : : : : : : : :
3040 TGCGGCTATATCGCGACATCACCGATGGGAAGATCGGCGCTCCACT 2991
132 ..GlyLeuLeuGlySerCysValGlnLeuMetValValGlyIleLeu 147
||||| : : : ||||| : : : : : : : : : : : : : : : : :
2990 TCGGCGCTATAGACGCTTGTTCGGCGGTATGTGTCGAGCGCCGCTG 2941
148 LeuAlaTyrLeuAlaGlyTyrVal...LeuGluTrpArgTrpLeuAlaVa 163
: : : ||| ||| : : : ||||| : : : : : : : : : : : :
2940 GCGGGGAGACTGTGGCGCGCATCTCTTCATGCACCATTCCTTGG... 2893
163 lleuGlyCysValProProSerLeuLeuLeuLeuMetCysPheMetP 180
: : : : : : : : : : : : : : : : : : : : : : : : : : :
2892 ...GCGGGGCTCTACAGGCTTCAACCTACTACTGCGCTGCTCTTA. 2848
180 roGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaIle 196
2847 ..... ATGAGAG..... 2839
197 AlaleuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspPro...Pr 212
```

```
2838 ..... ||::: || |
212 oIleGIyAla..... GluInSerPheHisLeuAlaLeuArgG 226
||::: || |
2809 CTTGAGACCTTCACACCAAGTCAGCTCTTCGGTGGGCG..... 2770
226 InProGIyIleTyrsProPheIleIleGIyValSerLeuMetAla... 241
::: :: || | :::: ||:::
2769 ..... CGGGGCACTACTATGCTGCCCACTTATGACTGTC 2734
242 ..... PheGIInLeuSerGIy... ValAsnAlaValMetPheTyraI 255
::: ||::: || | ||:::
2733 TTCTTATCATCACTCACTGAGACAGTCGCCGACGCTCTGCG..... 2689
255 aGIuThrIlePheGIuGIuAlaIalysPheIysAspSer..... 268
::: ||::: || | ::: ||:::
2688 ... GTCAATTTGGCGAGACCGCTTTCGTGAGCGCGACATGATGTCG 2643
269 .. LeuAlaSerValValGIyValIleGIuValIleuPheThrAla... 283
||::: ||::: ||::: ||::: ||:::
2642 GCCTTCGCTTCGGGATTCGGAATCTTCACGCCCTCGCTCAAGCCTTC 2593
||::: ||::: || | ||:::
284 ValAlaAlaLeuIleMetAspArgAlaGIyArgAlaLeuLeuVal... 299
||::: ||::: || | ||:::
2592 GTCACTGTCCCGCACCAACGTTTCGGCGAGAACAGCAGCATTAATCGC 2543
300 ..... LeuSerGIyValValMetValPheSerThr. 309
2542 CGGCATGGCGCGCGAGCGGCTGGGCTACGCTTCGCTTCGGCGGAGCG 2493
310 ..... SerAlaPheGIyAlaTyrsPheIysLeuThrgInGIyGIyPro 323
||::: ||::: ||::: ||::: ||:::
2492 GAGCGTGATGGCTTCCCATTAATGATTCCTCGCTTCGGCGGCGCATC 2443
324 GIyAsnSerSerHisValAlaIleSerAlaProValSerAlaGIInProVa 340
||::: ||::: || | ||:::
2442 GGGATGCC..... GCGTTGCAGGCCATGCTCTCCAGCGAGGTAGA 2402
340 lAspAlaSerValGIy..... LeuAlaTrpLeuAla.... 350
|| |
2401 TGACGACCATCAAGGACACCTTCACAGGATCGCTCGGCGCTTTACGAGCC 2352
351 ..... ValGIySerMetCysLeu..... PheIleAlaGIyPhe 361
::: ||::: ||::: ||::: ||:::
2351 TAACTTCGATCACTGGACCGGCTGATGTCACGGCATTTATGCCGCTCG 2302
362 AlaValGIyTTPGIyProIleProTrpLeuLeuMetSerGIuIlePhe 377
||::: ||::: ||::: ||::: ||:::
2301 GCGAGCACATGAGACGGGTTGGCATGATTTAGCGCGCGCCTATATAC 2254
```


160 TTPleuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMe 176
|||||
155 TGCGTGGCTGTGCTGGCTGGCCCTCCCTCATGTGCTCTCAT 204
176 TCSPheMetProGluThrProArpPheLeuLeuThrGlnHisArgArg 133
|||||
205 GTGCTTACAGCCCGAGACCCCGGCTTCCGTGCTGACTCAGCACAGCGCC 254
193 TngLualalle...AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrp 208
|||||
255 AGAGAGCCCATGGCGCCCTGGCGTCTCTGTGGGGCTCCGAGCAGCGCTGG 304
209 GluAspProProGluGlyAlaGlnSerPheHisLeuAlaLeuLeuAr 225
|||||
305 GAAGACCCCGCCCATCGGGCTGAGCAGAGCTTTCACCTGGCCCTCGCTGG 354
225 ggInProGlyIleTyrTrpSerProPheIleIleGlyValSerLeuMetLap 242
|||||
355 GCGAGCCGCGATCTACAGCCCTTTCATCATCGGCGTCTCCCTCATGGCCT 404
242 heGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIle 258
|||||
405 TCCAGCAGCTGTGGGGGTCAACGCCCTCATGTTCTATGACAGACCATC 454
259 PheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlu 275
|||||
455 TTTGAAGAGCGCCAACTTCAAGCAGACAGCGCTGGCTCGGCTCGTGGG 504
275 yValIleGluValLeuPheThrAlaValAlaAlaLeuIleMetAspArg 292
|||||
505 TGCATCCAGCGGTCTGTTCACAGCTGTGGCGGCTCTCATCATCAGACAG 554
292 laGlyArgArgLeuLeuLeuValLeuSerGlyValValMetValPheSer 308
|||||
555 CAGGGGAGAGCGCTGCTCTGCTTGTCAAGTGTGTCATGAGGTTCAGC 604
309 ThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyProGlyAs 325
|||||
605 AGCAGTGCCTTCGGCGCTACTTCAAGCTGACCCAGCGTGGCCCTGGCAA 654
325 nSerSerHisValAlaIleSerAlaProValSerAla.GlnProVal.As 341
|||||
655 CTCTCGCAGCGTGGCCATCTCGGCGCTCTCTGTCACAGCCCTGTGTA 704
341 pAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPhe 357
|||||
705 TCCACAGCGTGGGCTGCTGCTGGGCGCTGGGACAGCATGTGCTTCT 754
358 IleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSe 374
|||||
755 ATCGCGCTTTTGGCGTGGGGCTGGGCGCATCCCTTGGCTCTCATATGC 804
374 rgLutIlePheProLeuHisVal 381
|||||
805 AGAGATTTCCTCTGTGCTGC 826
seq_name: gb_est2:B1334832
seq_documentation_block: 991 bp mRNA EST 30-JUL-2001
LOCUS B1334832
DEFINITION 602998959P1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5141007 5',
mRNA sequence.
ACCESSION B1334832
VERSION B1334832
KEYWORDS B1334832.1 GI:15019489
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 991)
AUTHORS NIH-MGC <http://mgi.mgi.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/HLN at:
<http://image.llnl.gov>
plate: LLNL1347 row: f column: 16
High quality sequence strip: 835.
Location/Qualifiers
1..991
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5141007"
/clone_lib="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NOL1;
Site_2: Salt: Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

BASE COUNT 138 a 326 c 311 g 216 t
ORIGIN

alignment_scores:
Quality: 1233.50 Length: 311
Ratio: 4.328 Gaps: 11
Percent Similarity: 91.640 Percent Identity: 86.817

alignment_block:
US-09-516-493-7 x B1334832 ..

Align seq 1/1 to: B1334832 from: 1 to: 991

120 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuGlySe 136
|||||
73 GTCTACATCTCCGAATCGCTACCAGACATCCGGGGGTTCGCTGC 122
136 rCysValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaG 153
|||||
123 CTGTGTACAGCTAATGTGTGCTGCGCATCTCTGCGCTTACCTGCAG 172
153 TyrTrpValLeuGluThrArgTrpLeuAlaValLeuGlyCysValPro 169
|||||
173 GCTGGTGTGTGAGTGGCGCTGGCTGTGCTGGCTGCTGCCCTCC 222
170 SerLeuMetLeuLeuLeuMetCysPheMetProGluThrProArgPhe 186
|||||
223 TCCCTCATGTGCTTCTCATGTGCTTATGCGCCAGACCCCGGCTTCT 272
186 uLeuThrGlnHisArgArgGlnGluAlaIle...AlaLeuArgPheLeu 202
|||||
273 GCTGACTACAGACAGCGCCAGGACCATGGCGCCCTGCGTTCGT 322
202 rPglySerGluGlnGlyTrpGlnAspProProIleGlyAlaGluGlnSer 218
|||||
323 GGGGCTCCGAGCAGCGCTGGGAGAACCCCTCATGGGGCTGAGACAGGC 372
219 PheHisLeuAlaLeuLeuArgGlnProGlyIleTyrTrpSerProPheIle 235
|||||
373 TTTTACCTGGGCGCTGCGGCGACCGCGCATCTACAAAGCCCTTCATCAT 422
235 eGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValM 252
|||||
423 CCGGCTCTCCCTGATGGCTTCCAGCAGCTGTGGGGGTCAACCCGCA 472
252 eTyrTrpAlaGluThrIlePheGluGluAlaLysPheLysAspSerSer 268
|||||
473 TGTTCATGACAGACCATCTTTGAGAGGCCAAGTTCAAGAGACAGACG 522

```

269 LeuAlaSerValValValGlyValIleGlnValLeuPheThrAlaValAl 285
|||||
523 CTGGCCTCGGCGTGTGGGTGTCATCCAGTGGCTTCCACAGCTGGC 572
|||||
285 aAlaLeuIleMetAspArgAlaGlyArgValLeuLeuValLeuSerG 302
|||||
573 GGCTCTATCATCATGAGCAGACGAGCGAGGCTCTCTGCTGTGTGAG 622
|||||
302 lValValMetValPheSerThrSerAlaPheGlyAlaTyrPheValLeu 318
|||||
623 GTGGGTCATGCTGTTCAGCAGCAGTGGCTTGGCGCTTACTTCAAGCTG 672
|||||
319 ThrGlnGlyGlyProGlyAsnSerSer.HisValAlaIleSerAlaPro 335
|||||
673 ACCAGAGGTGGCCCTGGCAACTCTCGCAAGCTGGCCATCTGGCGGCTG 722
|||||
335 AlSerAlaGln.ProValAspAlaSerVal.GlyLeuAlaTyrPheVal 351
|||||
723 TCTCTGCAACAAGCTGTGTGATGCCAGCGTGGGCTGGCTGGCGCG 772
|||||
351 Al.GlySerMetCysLeuPheIleAlaGlyPheAlaValGly...TyrG 366
|||||
773 TTGGGCGAGCATGTGCTCTTCATGCGCGGCTTGGCGTGGGCTTGGG 822
|||||
366 TyrPheIleProTyrLeu...LeuMetSerGluIlePheProLeu...His 381
|||||
823 CCCATTCCCTCGGCTGCTCATTTGTGAGAAATCTTCCCTCTGCATATG 872
|||||
381 AllysGlyValAlaThrGlyIleCysValLeuThrAsnTyrPheMet 397
|||||
873 TCACGGGCGCTGGCAGCAGGCAATTCAGTCCCAACAATGGCTCCATG 922
|||||
397 lAbPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 413
|||||
923 GCCTTCCCGGAGACAGAGATTGACGCCACTTGGAAAGTCCCAAGGCC 972
|||||
414 TyrGlyAlaPheTyrPheAla 420
|||||
973 TTGAAC...CTCTGGTGTGCT 990
|||||
seq_name: gb_est2:BG478000

```

```

seq_documentation_block: 828 bp mRNA EST 21-MAR-2001
LOCUS BG478000 603522606F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4641145 5',
DEFINITION mRNA sequence.
ACCESSION BG478000.1 GI:13410279
VERSION BG478000.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 828)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rstraubs@nihs.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1001407 row: k column: 02
High quality sequence stop: 788.
Location/Qualifiers
1..828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4641145"

```

```

/clone.lib="NIH_MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCGCAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 85 a 309 c 275 g 159 t
ORIGIN

```

```

alignment_scores:
Quality: 1091.50 Length: 239
Ratio: 4.766 Gaps: 4
Percent Similarity: 95.816 Percent Identity: 94.142

```

```

alignment_block:
US-09-516-493-7 x BG478000 ..

```

```

Align seg 1/1 to: BG478000 from: 1 to: 828

```

```

1 ArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPh 17
|||||
116 CGCGCGGCTTCTCTCGCGGCTTGGCGCTTCCCTGGCGCCACTGAGCTT 165
|||||
17 eGlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArg 34
|||||
166 CGGCTTGGCGCTCGGTACAGCTCCCGCCATCCCTGACAGGCGG 215
|||||
34 lAlaIleProAlaProArgLeuAspAspAlaAlaIleSerTyrPheGly 50
|||||
216 CGCGCGCCCGCGCGCGCTGCGACGAGCGCGCGCTCTGCTGGTGGG 265
|||||
51 AlaValAlaThrLeuGlyAlaAlaAlaGlyValLeuGlyTyrPhe 67
|||||
266 GCTGCTGAGACCTGGGTGCGCGCGGAGGAGTGTGCGGCTGGCT 315
|||||
67 uValAspArgAlaGlyArgLysLeuSerLeuLeuCysSerValProp 84
|||||
316 GGTGAGCCCGCGCGCGGCAAGCTGAGCCTTCTGCTGCTGCGCT 365
|||||
84 heValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTyrPheLeu 100
|||||
366 TCGTGGCGCGCTTGGCCGTCATCACCGCGCGCGGACGTCGTGATCTG 415
|||||
101 LeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 117
|||||
416 CTGGGGGGCGCGCTCTCACCGGCGCTGCGCTGCTGCTGCTCTTCACT 465
|||||
117 lAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeu 134
|||||
466 GGCGCCCGGCTCATATCTCGGAATGCGCTTACCGCAGCGTGGGCTGG 515
|||||
134 euGlySerCysValGlnLeuMetValValValGlyIleLeuAla..T 150
|||||
516 TCGGCTCTCTGTGACAGCTAATGTCGTCTCGGCATCTCTTGGGCTT 565
|||||
150 TyrLeuAlaGlyTyrPvalLeuGluTyrPargTyrPheAlaValLeuGlyCys 166
|||||
566 ACCTGGAGAGGCTGGGTCTGAGATGGCGGCTGGCTGGCTGGCTGGC 615
|||||
167 Val.ProProSerLeuMetLeuLeuLeuMetCysPheMetProGluThrP 183
|||||
616 GTTGGCCCGCTTCCCTCATGCTGTCTTCATGTGCTTATCCCGAGAGCC 665
|||||
183 roArgPheLeuLeuThrGlnHisArgArgGlnGluAlaIle...AlaLeu 198
|||||
666 CGAGCTTCTGCTGACTGACACAGGCGCGAGAGGACCATGGCGGCTGC 715
|||||

```

```
199 ArgPheLeuTrpGlySerGlnGlnGlyTrpGluAspProProlIleGlyAl 215
      |||
      |||
716 GGTTCCCTGACGCTCCGACAGGCTGCGAGACGCCCCCATCGGGCG 765
215 aglGlnSerPheHisLeuAla..LeuLeuArgGlnProGlyIleTyrLy 231
      |||
      |||
766 TGAGAGAGCTTTCACCTTGGCCCTGTTGGCGAGAGCCCGCATCTACAA 815
231 sPropheIleIle 235
      |||
      |||
816 GCCCTTCATCATC 828
seq_name: gb_est2:BG7171736
seq_documentation_block: 879 bp mRNA EST 15-MAY-2001
LOCUS BG7171736
DEFINITION 602720404F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4837297 5',
      mRNA sequence.
ACCESSION BG7171736
VERSION BG7171736.1 GI:14082389
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 879)
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs.femail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
http://image.lnl.gov
Plate: LLM10769 row: h column: 02
High quality sequence stop: 765.
FEATURES
Location/Qualifiers
source 1..879
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4837297"
/clone_1ib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (glcgaq
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.2 kb and
normalized to Rot 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 95 a 319 c 296 g 169 t
ORIGIN
```

alignment_scores:

```
Quality: 1091.50 Length: 261
Ratio: 4.419 Gaps: 10
Percent Similarity: 94.636 Percent Identity: 91.571
```

alignment_block:

US-09-516-493-7 x BG7171736 ..

Align seg 1/1 to: BG7171736 from: 1 to: 879

```
1 ArgArgValPheLeuAlaLeuPheAlaLeuGlyProLeuSerPh 17
|||||
```

```
105 CCGCCGCTTCTTCCTCGCCGCTTCGCCGCTGCCCTGGGCCCACTAGCTT 154
17 eGlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArg 34
      |||
      |||
155 GCGCTTCGCGCTCGGCTACAGCTCCCGGCCATCCTACCTGCACCGCG 204
34 lAlaIleProAlaProAlaProAlaGluAspAspAlaAlaIleSerTrpPheGly 50
      |||
      |||
205 CCGCGCCCGCCCGCCCGCCCGCTGGACGACGCGCGCCCTCTCGTTGCGG 254
51 AlaValAlaThrLeuGlyAlaAlaAlaGlyGlyValLeuGlyGlyTyrPhe 67
      |||
      |||
255 GCTGTCGTGACCCGTGGTCCCGCGGGGAGAGTCTGGGCGGTGGCT 304
67 uValAspArgAlaGlyArgLysLeuSerLeuLeuLeuGlyCysSerValProp 84
      |||
      |||
305 GGTGGACCCCGCCCGCGCGGCAAGCTGAGCTTTCCTGCTCTCCGTGCCCT 354
84 heValAlaGlyPheAlaValAlaThrAlaAlaGlnAspValTrpMetLeu 100
      |||
      |||
355 TCGTGCGCGGCTTGGCGCTCATACACCGCGCCGACGCTGGATGCTG 404
101 LeuGlyGlyArgLeuLeuThrGlyLeuAlaGlyGlyValAlaSerLeuVal 117
      |||
      |||
405 CTGGGGGGCGGCGCTCTCACCGCGCTGGCGCTGGGCTGGCTCTCTAGT 454
117 lAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeu 134
      |||
      |||
455 GCGCCCGGCTTCATCTCCGAAATCGCCTACCCAGACAGTCCGGGGGTTCG 504
134 euGlySerCysValGlnLeuMetValAlaValGlyIleLeuLeuAlaTyr 150
      |||
      |||
505 TCGGCTCCTGTGTGACAGTAATAGTGTGTCGTGCGCATCTCTCGGCTAC 554
151 LeuAlaGlyTrpValLeuGluTrpArgTrpLeuAlaValLeuGlyCysVal 167
      |||
      |||
555 CTGGCAGGCTGGGTGCTGGAGTGGCGCTGGCTGGCTGGCTGGCTGGCT 604
167 lProProSerLeuMetLeuLeuLeuMetCysPheMetProGluThrProAla 184
      |||
      |||
605 GCGCCCGCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 654
184 rGpPheLeuLeuThrGlnHisArg.ArgGlnGluAlaIleAla..LeuArg 199
      |||
      |||
655 GCTTCTCTGTGACTGACATGCGACATGCGCGCAGAGGACATGGCTGCGCGG 704
200 PheLeuTrpGlySerGlnGlnGlyTrpGluAspProProlIleGlyAlaG 216
      |||
      |||
705 TTCTCTGTGGGCTTCGAGAGGCTGGGAAGAACCCCATCTGGGGCTGA 754
216 u.GlnSerPheHisLeu..AlaLeuLeuArg.GlnProGlyIle.TyrLys 231
      |||
      |||
755 GCCAGAGCTTTCACCTCGGCGCTGTGGCGACAGCCCGGATCTTACAGA 804
232 .sPropheIleIle.GlyAlaSerLeuMet.AlalaPheGlnLeuSerG 247
      |||
      |||
805 GCCCTTCATCATCGGCGCTTCCCTGATGAGCTT...CAGCAGCTGCTGG 851
247 yValAsnAlaValMet 252
      |||
      |||
852 GGTCAAGGCGCGCATG 867
```

seq_name: gb_est2:BG717034

seq_documentation_block: 777 bp mRNA EST

LOCUS BG717034 602689093F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4821398 5',

DEFINITION mRNA sequence.

ACCESSION BG717034

VERSION BG717034.1 GI:13996221

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 777)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroaki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM10728 row: 1 column: 23
 High quality sequence stop: 722.

FEATURES
 source
 1..777
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4821598"
 /clone_1ib="NIH-MGC_97"
 /lab_host="DH10B"
 /note="Organ: testis; Vector: pBluescript (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.2 kb and
 normalized to 50x. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI, National
 Institutes of Health). Note: this is a NIH-MGC Library."

BASE COUNT 76 a 295 c 263 g 143 t
 ORIGIN

alignment_scores:

Quality: 1086.50 Length: 225
 Ratio: 4.916 Gaps: 2
 Percent Similarity: 98.222 Percent Identity: 96.889

alignment_block:
 US-09-516-493-7 x BG717034

Align seg 1/1 to: BG717034 from: 1 to: 777

```

1  ArgatgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPh 17
   |||||||
105  CCGCGCTTCCTCGCGCGCTTCGCGCGCTGCGCCACCTGACCTG 154
   |||||||
17  eGlyPheAlaLeuGlyTyrSerSerProAlaLeuProSerLeuGlnArg 34
   |||||||
155  CGGCTTCGCGCTGCTACAGCTCCCGCGCCATCCCTGACCTGACGCG 204
   |||||||
34  laAlaProProAlaProArgLeuAspAspAlaAlaSerTrpPheGly 50
   |||||||
205  CCGCGCGCGCGCGCGCGCTGAGACGCGCGCTTCCTGCTGCTGCTG 254
   |||||||
51  AlavaValThrLeuGlyAlaAlaAlaGlyValLeuGlyGlyTrpLe 67
   |||||||
255  GCTGCTGACCTGCGGTGCGCGCGCGGAGTGTGCGGCGCTGCTGCT 304
   |||||||
67  uValAspArgAlaGlyArgLysLeuSerLeuLeuLeuGlySerValProp 84
   |||||||
305  GGTGACCGCGCGCGCGCGCAAGCTGAGCTTGTGCTGCTGCGGCTG 354
   |||||||
84  heValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeu 100
   |||||||
355  TCGTGGCGCGCTTGGCTGATCACCAGCGCGCGGAGAGTGTGAGTGTG 404
   |||||||
101  LeuGlyGlyArgLeuLeuThrGlyLeuAlaGlyGlyValAlaSerLeuVa 117

```

```

|||||
405  CTGGGGGGCGCGCTTCCTACCGCGCTGCGCTGCTGCTGCTGCTAGT 454
|||||
117  lAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuL 134
|||||
455  GCGCGCGCTTACATCTCCGAATGCGCTACCGACAGCTCCGGGGCTTGC 504
|||||
134  euGlySerCysValGlnLeuMetValValAlaGlyIleLeuLeuAlaTyr 150
|||||
505  TCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 554
|||||
151  LeuAlaGlyTrpValLeuGluTrpArgTrpPheAlaValLeuGlyCysVa 167
|||||
555  CTGGCAGGCTGGTGTGCTGAGTGGCGGCTGCTGCTGCTGCTGCTGCTG 604
|||||
167  lProProSerLeuMetLeuLeuLeuMetCysPheMetProGluTrpProA 184
|||||
605  GCGCGCGCTCCCTCATGCTGCTTCATGCTGCTGCTGCTGCTGCTGCTG 654
|||||
184  rPheLeuLeuThrGlnHisArgArgGlnGluAlaIle..AlaLeuArg 199
|||||
655  GCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 704
|||||
200  PheLeu.TrpGlySerGlnGlnGlyTrpGluAspProProIleGlyAlaG 216
|||||
705  TTCCTGTTGGGGCTCCGAGCATGCTGGAAGACCCCGCCATC.GGACTG 753
|||||
216  lGlnSerPheHisLeuAlaLeu 223
|||||
754  ACCAGAGCTTCACCTGACCTG 776
|||||
seq_name: gb_est2:BG281777
seq_documentation_block: 838 bp mRNA EST 21-FEB-2001
LOCUS BG281777
DEFINITION 602402292P1 NIH-MGC_20 Homo sapiens cDNA clone IMAGE:4544686 5',
mRNA sequence.
ACCESSION BG281777
VERSION BG281777.1 GI:13030703
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 838)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DC70/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
Plate: LLM1228 row: 9 column: 23
High quality sequence stop: 784.
Location/Qualifiers
1..838
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4544686"
/clone_1ib="NIH-MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOT7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb) Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit

```

FEATURES

source
 1..838
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4544686"
 /clone_1ib="NIH-MGC_20"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOT7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb) Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit


```

244 ..... 244
67 uvalasparagialglyargylsleuSerleuLeuLeuCysSerValProp 84
244 ..... 244
84 hevalalaglyphealavalallethralalaglinaspValITrPmetLeu 100
244 ..... 244
101 leuglyglyargyleuLeuThrglyLeuAlaCysglyValAlaSerLeuVa 117
244 ..... 244
117 lalaprovalITyrIleSerGluIleAlaTyrProAlaValArgglyLeuL 134
245 .....GTCACATCTCCGAATCCCTACACGACGACGACGCGGGGCTTGC 287
134 euglySerCysValGlnLeuMetValValValglyIleleuLeuAlaTyr 150
288 TCGGCTCCTGCTGCGACGCTAATGCTGCTGCGCATCTCTGCGCTTAC 337
151 leualaglyTrpValLeuGluTrpArgTrpLeuAlaValLeuGlyCysVa 167
338 CTGGCAGCGCTGGTCTGTGAGTGGCGCTGGCTGGCTGGCTGGCT 387
167 lproproSerleuMetleuLeuLeuMetCysPheMetProGluTrhProA 184
388 GCCCCCTCCCTCATGCTGCTTCATGCTTCATGCTTCATGCTTCATGCT 437
184 rgPheLeuLeuThrglnHisArgArgGlnGlnAlaIle...AlaLeuArg 199
438 GCTTCCTGCTGCTGCTGACGACGACGCGCCAGAGGCGCATGGCGCGCG 487
200 PheLeuTrpGlySerGlnGlnGlyTrpGluAspProProIleGlyAlaG1 216
488 TTCTGCTGGGCTCGAGCAGGCGCTGGGAAGACCCCATCGGGGCTGA 537
216 uGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProp 233
538 GCACAGCTTTCACCTGGCGCTGCTGGCGACGCGCGCATCTACAGACCT 587
233 heileileglyValSerleuMetAlaPheGlnGlnLeuSerGlyValAsn 249
588 TCATCATGCGGCTGCTCCCTGATGGCTTCCAGACACTGTGGGGTCAAC 637
250 AlalalmetPheTyrAlaGluTrhIlePheGlnGlnAlaLysPheLys.. 265
638 GCCGTGATGTTCTATGCAGAGACCATCTTTGAAGAGGCCAAGTTCAAGCG 687
266 AspSerSerleuAlaSerValValValglyValIleGlnValLeuPheTh 282
688 ACCGACGACGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGACC 737
282 rAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuV 299
738 AGCTGTGGCGCTCTCATATTGCACAGACGACGCGGAGGCTGCTGCTGG 787
299 al 299
788 TC 789
seq_name: gb_est2:BE910478

```

```

seq documentation block:
LOCUS BE910478 934 bp mRNA EST 20-OCT-2000
DEFINITION 601501005F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3903048 5',
ACCESION BE910478
VERSION BE910478.1 GI:10407108
KEYWORDS EST.
SOURCE human.

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 934)
AUTHORS NIH-MGC. http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLN at:
http://image.lln.gov
Plate: LHM9706 row: 1 column: 01
High quality sequence stop: 710.
FEATURES
source
location/Qualifiers
1..934
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3903048"
/clone_lib="NIH_MGC_70"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pCMV-SPORT6. Site1: NotI;
Site2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 Kb. Library constructed by Life
Technologies."
BASE COUNT 143 a 291 c 287 g 213 t
ORIGIN

```

```

alignment_scores:
Quality: 1036.50 Length: 238
Ratio: 4.566 Gaps: 2
Percent Similarity: 95.378 Percent Identity: 93.277

```

```

alignment_block:
US-09-516-493-7 x BE910478

```

```

Align seg 1/1 to: BE910478 from: 1 to: 934

```

```

217 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProp 233
13 AGACCTGTCACCTGGCGCTGCTGGCGACGCGCGCATCTACAGCCCTT 62
233 eIleIleGlyValSerleuMetAlaPheGlnGlnLeuSerGlyValAsn 250
63 CATCATGGGTCTCTCCCTGATGGCTTCCAGCAGCTGTGGGGTCAACG 112
250 lAlaIleMetPheTyrAlaGluTrhIlePheGlnGlnAlaLysPheLysasp 266
113 CCGTCATGTTCTATGCAGACCATCTTTGAAGAGGCCAAGTTCAAGAGC 162
267 SerSerleuAlaSerValValValglyValIleGlnValLeuPheThrAl 283
163 AGCAGCCTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGACC 212
283 aValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuValL 300
213 TGTGGCGGCTTCTCATGACAGACGACGAGGCGGAGGCTGCTGCTGCTGCT 262
300 euSerGlyValAlaMetValPheSerTrpSerAlaPheGlyAlaTyrPhe 316
263 TGTGAGGTGTGTATGATGTTTTCACGACGAGTGCCTTGGGCGCTACTTC 312
317 LysLeuThrglnGlyGlyProGlyAsnSerSerHisValAlaIleSerAl 333
313 AAGCTGACCCA.GGTGGCGCTGGACACTCTCTGCGACGTCGCGCATCTGCGC 361
333 aprovalSerAlaGlnProValAlaAspAlaSerValGlyLeuAlaTrpLeuA 350

```


Seq primer: -21M13 Reverse
High quality sequence stop: 595
POLYA-No.

FEATURES

source

Location/Qualifiers
1. 595

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="H3037C03"
/clone="H3037C03"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"

/note="Vector: pSPORT1. Site_1: SalI; Site_2: NotI. This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

BASE COUNT 95 a 177 c 182 g 141 t
ORIGIN

alignment_scores: Quality: 847.00 Length: 194
Ratio: 4.458 Gaps: 2
Percent Similarity: 97.938 Percent Identity: 84.021

alignment_block:
US-09-516-493-7 x BG079217 ..

Align seg 1/1 to: BG079217 from: 1 to: 595

```

197 AAlaLeuArpHeLeuTrpGlySerGlnGlnGlyTrpGluAspProPro1 213
|||||
15 GCGTTGGCGCTTCCTGCTGGCGCTTCAGAGCGCTCGGAAGACCCCTGT 64
|||||
213 eglValaGlu...GlnSerPheHisLeuAlaLeuLeuArpGlnProGly1 225
|||||
65 TGGGGCTGACGACCAAGCGCTTCACACTGCCCTGCTGAGGCCCTTGGA 114
|||||
229 leTyrLysProPhe1le1leGlyValSerLeuMetAlaPheGlnGlnLeu 245
|||||
115 TCTCAAGACCCCTCATCTCGGCAATTCCTTCATGCTTCCTCAGACGCT 164
|||||
246 SerGlyValaAsnAlaValaMetPheTyrAlaGlnTrp1lePheGlnGlnVal 262
|||||
165 TCGAGGGCTCAATGCTATCATGTTCTATGCCAACACACATCTTCGAGGAGGC 214
|||||
262 alyPheLysAspSerSerLeuAlaSerValValValGlyVala1leGlnV 275
|||||
215 CAACTTCACAGACACAGACCGCTCGCTCGCTCAGCTGCGCATATCCAGG 264
|||||
279 aLeuPheThrAlaVala1AlaLeu1leMetAspArgAlaGlyArgArg 295
|||||
265 TCCTGTTCACTGCTGTGGCGGCCCTCATCATGACAGAGCGGGAAGG 314
|||||
296 leuLeuLeuValaLeuSerGlyVala1MetValaPheSerThrSerAlaPh 312
|||||

```

```

315 CTGCTCTCTGGCCCTTGCCGGCTGTGATCATGCTTCTAGTATGATGCCCTT 364
|||||
312 eglValaTyrPheLysLeuThrGlnGlnGlyProGlyAsnSerSerHisV 329
|||||
365 TGCTACCTACTTCAAACTGACAGAGCGCTCCCGACAGCACTCTCCACAG 414
|||||
329 a1Ala1leSerAlaProValSerAlaGlnProVala1spalaserValGly 345
|||||
415 TGAGGCTG...GTGCCATTCGCGCGAGCGCTGTGATGTCCCAAGTGGA 461
|||||
346 leuAla1TrpLeuAlaValaGlySerMetCysLeuPhe1leAlaGlyPheAl 362
|||||
462 CTGGCGCTGCTGCTGAGCAGCATGTGCTTCTTATGTCTGCTGCTTTCG 511
|||||
362 aValaGlyTrpGlyPro1leProTrpLeuLeuMetSerGln1lePhePro1 379
|||||
512 GCGGGCTGCGGACCATCCCTCGCTCTCATGTCTAGAGATCTTCCCTC 561
|||||
379 euHisValaLysGlyVala1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1 389
|||||
562 TGCATGTCAAGGCTGTGCTACCGCATCTGT 593
|||||

```

seq_name: gb_est2:B1402088

seq_documentation_block:

LOCUS B1402088 594 bp mRNA

DEFINITION M1-P-CP0-nvv-e-05-0-UI.s1 M1-P-CP0 Sus scrofa cDNA clone

ACCESSION B1402088

VERSION B1402088.1 GI:15181149

KEYWORDS EST.

SOURCE Sus scrofa

ORIGIN pig.

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Cetartiodactyla: Suidae: Sus.

REFERENCE 1 (bases 1 to 594)

AUTHORS Bernaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT

Contact: Tugle CK
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kildee Hall, Ames, IA 50011-3150, USA
Tel: 5152944252
Fax: 5152942401
Email: ckugle@iastate.edu

Oligo-dT track not found, Not 1 site shown in beginning of sequence is likely internal to the message. cDNA Library Preparation: M.B. Soares lab, University of Iowa EST sequencing: M.B. Soares lab, University of Iowa Clone distribution: clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward

POLYA-No.

FEATURES

source

Location/Qualifiers
1. 594

/organism="Sus scrofa"

/db_xref="taxon:9823"

/clone="M1-P-CP0-nvv-e-05-0-UI"

/clone="M1-P-CP0"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pRT3D-Pac (Pharmacia) with a modified

polylinker. Site_1: Not I; Site_2: EcoRI. The M1-P-CP0

library is derived from uteruses. For a detailed description

of the library from which this clone was derived, please

visit our web site at <http://pigst.genome.iastate.edu/>.

The procedure used to create this library has been

previously described (Bernaldo, Lennon and Soares, Genome

Research 6:791-806, 1996)

TAG_SEQ=None found"

BASE COUNT 58 a 221 c 192 g 121 t 2 others
ORIGIN


```

|||||
287 TCGGCTCCTGTCGAGCACTAATGTCGTCGGCATCTCCGCGCTAC 336
151 leuAlaGlyTrpValIleuGlnTrpArgTrpLeuAlaValIleuGlycysVa 167
|||||
337 CTGGGAGGCTGGGCTGCTGAGTGGCGCTGGCTGGCTGGCTGGCTGGCT 386
167 lProProSerLeuMetLeuLeuIleuMetCysPheMetProGlnThrProA 184
|||||
387 GCGCCCTCCCTCATGCTGCTTCATGCTTCATGCTTCATGCTTCATGCTTC 436
184 rghPheLeuLeuThrGlnHisArgArgGlnGlnAla...lIeAlaLeuArgP 200
437 GCTTCCTCTGCTGACTCAGCAGCAGCGCGCAGAGCGCATGTGCCCTCCGT 486
200 heLeuTrpGlySerGlnGlnGlyTrpGlnAspProGlnIleGlyAlaGln 216
487 TCCTTGCGGCTCCGACGAGCGCTGGAGACCCCGCCATCGGGGCTGAG 536
217 GlnSerPheHisIleuAlaLeuLeuArgGlnProGlylIeTrpLysProh 233
537 CAGACCTTTCACCTGCGCCCTGCTGGGAGC...CGGATCTACAGCCCTT 583
233 eIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsn 249
584 CATCATCGGGTTC.TCCCTCATGCGCTTCAGCAGAGCTGCGGGGTCAAC 632
250 AlaValIleMetPheTrpAlaGlnTrpIlePheGlnGlnAlaValIlePheLys 265
633 GCGGCTATGTTCTTATGACAGACATCTTTGAAGAAGC.AAGTTCAAGA 681
266 pSerSerLeuAlaSerValValAlcValIleGlnValLeuPheThrA 283
682 CAGCAGNCCTGGC.TCGGTGTCGGGGCTGTTCATCCAGTGCCTC.AAAG 729
283 lAlaValAlaIleuIleMetAspArgAlaGlyArgArgLeuLeuVal 299
730 CTGTGGGGGTCTCATTCAGACGCGAGGAGGAGGATGGCCCTGGCT 779
300 LeuSerGlyValIleMetValPheSerThrSerAlaPheGly..... 313
780 GGAAGGGG.....GCATGGGGCCACAGCCGA 805
313 ..... 313
806 ATGCTTCGGGGCAATTAAATGACCAAGCGGGGCAATCTCCCAAGG 855
314 .....AlaTrpPheLysLeuThrGlnGlyGlyProGlyAsnSerSer 327
856 GCATTGCGGGGCTGTAACGCTTAACCCGGGGTGGCGGGGCGACAGGAAGC 905

seq_name: gb_est2:BF140667
seq_documentation_block:
LOCUS      BF140667      917 bp      mRNA      EST      24-OCT-2000
DEFINITION 601786917f1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4014605 5',
            mRNA sequence.
ACCESSION  BF140667
VERSION     BF140667.1  GI:10979707
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgaabs-remail.nih.gov
COMMENT     Tissue Procurement: Gilbert Smith, Ph.D.
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

```

```

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LHAM9259 row: a column: 06
High quality sequence stop: 605.

FEATURES
    source
        1..917
            /organism="Mus musculus"
            /strain="C57BL/6"
            /db_xref="taxon:10090"
            /clone="IMAGE:4014605"
            /issue_type="NCI_CGAP_Lu30"
            /tissue_type="tumor, metastatic to mammary"
            /lab_host="DH10B"
            /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; transgenic model MMTV-1, expression driven by
            MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
            dt. Library constructed by Life Technologies.
            Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT      163 a      263 c      277 g      214 t
ORIGIN

alignment_scores:
    Quality:      830.00      Length:      274
    Ratio:        3.458      Gaps:      10
    Percent Similarity: 87.591      Percent Identity: 69.343

alignment_block:
US-09-516-493-7 x BF140667 ..

Align seg 1/1 to: BF140667 from: 1 to: 917

197 AlaLeuArgPheLeuTrpGlySerGlnGlnGlyTrpGlnAspProGln 213
|||||
9 GCCCTGCGCTTCTGTCGGGGCTCTGAGAGGCTGGAGAGAGCCCTGT 58
213 eGlyAlaGlnGln..SerPheHisIleuAlaLeuLeuArgGlnProGlyl 229
|||||
59 TGGGGCTGAGACACAGGCTTCCAGCTGGCCCTGAGGGCGCTGGCAT 108
229 eTrpLysProPheIleIleGlyValSerLeuMetAlaPheGlnLeuLys 246
|||||
109 CTCAAGCCCTCATCATCGCATTCCTCATGCTTCACAGAGCTGT 158
246 eGlyValAsnAlaValMetPheTrpAlaGlnTrpIlePheGlnGlnAla 262
|||||
159 CAGGGGTCAATGCTATCATCTTCTATGCCAAGACATCTTCAGAGAGGC 208
263 LysPheLysAspSerSerLeuAlaSerValValValGlnValIleGlnVal 279
|||||
209 AAGTTCAGAGCAGACAGCCTGGCTCGCTGCTGCTGCGCATTAATCCAGGT 258
279 lLeuPheThrAlaValAlaIleuIleMetAspArgAlaGlyArgArgL 296
|||||
259 CCGTTCACAGCTGTGGCGGC.CTCATCATGACAGCAGCGGCGGAAGGC 307
296 eLeuLeuValIleuSerGlyValValMetValPheSerThrSerAlaPhe 312
|||||
308 TGCTCTGGCCCTTGTGCGGTGATGATGCTTCATGATGATGATGCTT 357
313 GAlaAlaTrpPheLysLeuThrGlnGlnGlyProGlyAsnSerSerHisVa 329
|||||
358 GGTACCTACTTCAACATGACCGAGCGCTCCACAGCACTCTCCACGT 407
329 lAlaIleSerAlaProValSerAlaGlnProValAspAlaSerValGlyL 346
|||||
408 AGGCTGT...GTGCCCATCGCGGCGAGCCTGTGATGATGATGATGAG 454
346 eValTrpPheLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAla 362
|||||
455 TGGCTTGGCTGGCTGTAGGACATGTGCTTCATTCATTCGCTGGCTTGG 504

```

```
363 ValGlyTrpGlyProIleProThrLeuLeuMetSerGluIlePheProLe 379
|||||
505 GTGGGCTGGGACCTATCCCTGGCT.CTCATGTCAGAGATCTTCCCTCT 553
|||||
379 uHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuM 396
|||||
554 GCATGTCAAGGGGTGGCTACCGGCGTCTGTGTCTCACCACACTGGTTCA 603
|||||
396 eAlaPheLeuValThr.....LysGluPhe.....SerSerLeuMet 408
|||||
604 TGGTCTTTTCTTAACTTGACCAAAAGAGTTCAAGGCAGGTCACAAAGAA 653
|||||
409 GluValLeuArg.ProTyrGlyAlaPheTrpLeuAlaSerAlaPheCysI 425
|||
654 GATGGCTCAAGAACCGTTAACGGGTGGCTTACTGGGTTCCACCGGTTGGG 703
|||
425 le.....PheSerValLeuPheThrLeuPhe 433
|||
704 TTTCTTGGGACACCTCAAGGTGGTCCCATTTACACATGGAGCAGTTG... 750
|||||
434 CysValProGlu.ThrLysGlyLys..ThrLeu...GluGlnIleThrAla 448
|||||
751 ...GTCCCTGAGGACTAAAGCGCGGATCTGTGGACACAGTCTCCACGGC 797
|||||
449 HisPheGluGlyArg 453
|||||
798 CACTTTCGAGGGAGA 812
```


1298 TCTGCTGACCAAGAGTTTAACACATCATGAGATCCTCAGACCCACG 1347
 102 lylaPheTrpleuThralAlaPheCysIleuSerValIeuPheThr 118
 1348 GCGGCTTCTGCTCACCCTGCTCTGTATCTCAGCGCTCTTTACG 1397
 119 LeuThrPheValProGluThrIlySglYarGThrIeuGluGlnIleThrAl 135
 1398 CTCACCTTTCTCCCTGAGACTAAAGCAGACCTCTGGACAAATACACAGC 1447
 135 ahISleu.ArgAspGlyAspGlyProIeuSerValIThrGlySerProGlu 151
 1448 CCAATTTCAGAGGAGGTGACGGACCTT.TCTGTGACTGGCAGCCCTGAG 1496
 152 LeuSerTrpleuArgValSerIlySglYValGluTrpProGln 165
 1497 CTGAGCTGGCTCGGCTTTCAAAAGAGTGGAGTGCCCTCAG 1538
 seq_name: gb_ro:RNO245935

seq_documentation_block:
 LOCUS RNO245935 2087 bp mRNA ROD 18-FEB-2000
 DEFINITION Rattus norvegicus mRNA for glucose transporter (GLUTX1 gene).
 ACCESSION AJ245935
 VERSION AJ245935.1 GI:7018604
 KEYWORDS glucose transporter; GLUTX1 gene.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 2087)
 AUTHORS Ibberson,M., Uldry,M. and Thorens,B.
 TITLE GLUTX1, a novel mammalian glucose transporter expressed in the
 central nervous system and insulin-sensitive tissues
 JOURNAL U. Biol. Chem. 275 (7), 4607-4612 (2000)
 MEDLINE 20138191
 REFERENCE 2 (bases 1 to 2087)
 AUTHORS Ibberson,M.R.
 TITLE Direct Submission
 JOURNAL Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
 and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
 CH-1005, SWITZERLAND

FEATURES
 source Location/Qualifiers
 1..2087
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 gene 30..1466
 /gene="GLUTX1"
 CDS 30..1466
 /gene="GLUTX1"
 /codon_start=1
 /product="glucose transporter"
 /protein_id="CA57529.1"
 /db_xref="GI:7018605"
 /translation="MSPEDPQETOPFLRSPGARPGRRVFLATFAALGPLSFQHAL
 GYSSPAIPRLRATPRLGDTAASMPAAVVLGAAGVGLGMLDRAGRLLSLIL
 CTVPFTEGAVITAARDVWLILGGRLTGAGVAVLAPVYSEIAYPAVRLLSGC
 VOLAVTGLIAYVAGVLEWRILAVLIGVPTMLILCMYMEPTRFILTOHOGEA
 MAALRFKSGSEGEPEPPYGAERHGFOLAMLRPGYHKPLIITICIMVTOQLSVNAI
 MEYANTFEERAFKDSLSAVVGIIOVLETAVALIMDRAGKLLILASGVIVFSM
 SAFGYFKLTQSGPSNSHVLGVLPIASAPADVHGLIAMLVGSMCLFLAGFAVMGP
 IPMLMSIEIFPLHIKGVATGCVLTNMFMAFLVTKDFNSIMETILRPYGAFWLTAFCI
 LSVLFTIIVPPTKRGRTLEQITAHPEGR"
 BASE COUNT 355 a 673 c 580 g 479 t
 ORIGIN

alignment_scores:
 Quality: 741.00 Length: 148
 Ratio: 5.041 Gaps: 1
 Percent Similarity: 99.324 Percent Identity: 97.973

alignment_block:
 US-09-516-493-10 x RNO245935 ..
 Align seg 1/L1 to: RNO245935 from: 1 to: 2087

19 LeuValProIleSerAlaGluProAlaIAspValHisIleuGlyLeuAlaTr 35
 1098 CTGGGCCCATCTCCGCGAGACCTGTGATGTTCACCGGGGCTGGCTTG 1147
 35 pleuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyT 52
 1148 CTGGCTCTAGGACACATGTGCTCTTTCATGCTGCTTTTGCAGAGGCT 1197
 52 rpgIlyProIleProTrpleuIeuMetSerGluIlePheProIeuHisIle 68
 1198 GGGAGCCCATCCCTCGGCTCCTCATGTCAGAGATCTTCCTTGACATC 1247
 69 lylSglYValAlaThrGlyValAlcYsValIeuThrasnTrpIleMetAlaPh 85
 1248 AAGGCTGGCTACCGGCGTGTGTCTCTCACCAACTGCTCATGCTCTT 1297
 85 eleuValThrIlySgluPheAsnSerIleMetGluIleLeuArpProTyrg 102
 1298 TCTGTGACCAAGAGTTTAACAGCATCATGAGATCTCAGACCCCTACG 1347
 102 lylaPheTrpleuThralAlaPheCysIleuSerValIeuPheThr 118
 1348 GCGGCTTCTGCTCACCCTGCTCTGTATCTCAGCGCTCTTTTACG 1397
 119 LeuThrPheValProGluThrIlySglYarGThrIeuGluGlnIleThrAl 135
 1398 CTCACCTTTCTCCCTGAGACTAAAGCAGACCTCTGGACAAATACACAGC 1447
 135 ahISleu.ArgAspGlyAspGlyProIeuSerValIThrGlySerProGlu 151
 1448 CCAATTTCAGAGGAGGTGACGGACCTT.TCTGTGACTGGCAGCCCTGAG 1496
 152 LeuSerTrpleuArgValSerIlySglYValGluTrpProGln 165
 1497 CTGAGCTGGCTCGGCTTTCAAAAGAGTGGAGTGCCCTCAG 1538
 seq_name: gb_ro:AB033418

seq_documentation_block:
 LOCUS AB033418 2189 bp mRNA ROD 11-APR-2000
 DEFINITION Rattus norvegicus glut8 mRNA for glucose transporter 8, complete
 cds.
 ACCESSION AB033418
 VERSION AB033418.1 GI:7592743
 KEYWORDS glucose transporter 8.
 SOURCE Rattus norvegicus CDNA to mRNA, clone.lib: testis.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 2189)
 AUTHORS Ishibashi,K.
 TITLE Molecular cloning of a new putative glucose transporter
 JOURNAL Published only in Database (2000) In press
 REFERENCE 2 (bases 1 to 2189)
 AUTHORS Ishibashi,K.
 TITLE Direct Submission
 JOURNAL Submitted (08-OCT-1999) to the DDBJ/EMBL/GenBank databases. Kenichi
 Ishibashi, Jichi Medical School, Pharmacology; Minami-Kawachi,
 Kawachi, Tochigi 329-0498, Japan (E-mail: kishibashi.jichi.ac.jp,
 Tel:81-285-58-7326, Fax:81-285-44-5541)
 FEATURES
 source Location/Qualifiers
 1..2189
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 /clone_lib="testis"
 gene 109..1548
 /gene="glut8"

CDS 109..1548
/gene="glut8"
/note="glut8 has 12 transmembrane segments"
/codon_start=1
/product="glucose transporter 8"
/protein_id="BA094383.1"
/db_xref="GI:759274"
/translation="MSPEDPOETOPLLRSPGARARGRVLATFAALGLPLSGFAL
GYSSPAIPSLRRTAPPALRLGDPAASWFGAVYTLGAAGVILGGMLDRAGRLSL
LCTVPTGFAVITTAARDVMMILGGRLTLGACVAVSLVAPVTSFIAPVAVRGLGSC
CVOLMVTGILLAVAGVLEMRRLAVLGCVPPTMLLMCMPTPRFLLOHOYOE
AMALRFLMGSEGEHPVGAEGHGFOLAMLRRCVHKRPLITIGLWFOQSGVNA
IMFANTLFEAKFKDSSLASVTGIIQVLFVAVALIMDRAGRLKLLALSGVIMVFS
MSAFGTYFKLTQSGPNSHVLVPIIAEPADVHLIMDLAVGSMCLFIAGFVGMG
PIPLMSEIFPLHIKGVATGCVLTNNMFALVTKEFNSIMEILRPYGAFLMTAAFC
ILSVFTLTVPETKGRITLQVTAHFEGR"

BASE COUNT 392 a 702 c 609 g 486 t

ORIGIN

alignment_scores:
Quality: 741.00 Length: 148
Ratio: 5.041 Gaps: 1
Percent Similarity: 99.324 Percent Identity: 97.973

alignment_block:
US-09-516-493-10 x AB033418 ..

Align seg 1/1 to: AB033418 from: 1 to: 2189

19 LeuValProIleSerAlaGluProAlaAspValHisLeuGlyLeuAlaTr 35
|||||
1180 CTGGTGGCCATCTCCGAGAGCCGTGATGTTCCACCTGGGCGTGGCCTG 1229
35 pLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTr 52
|||||
1230 GCTGGCTGTAGGAGCAATGTCCTTCATCCGCTTTCAGTTCAGGCT 1279
52 rPGLyProIleProTrpLeuLeuMetSerGluIlePheProLeuHisIle 68
|||||
1280 GGGGACCCATCCCTGGCTGCTCATGTCAAGATCTTCCCTGTCACATC 1329
69 LysGlyValAlaThrGlyValCysValLeuThrAsnTrpPheMetAlaPhe 85
|||||
1330 AAGGCTGTGGCTACCGCGCTGTCTGTCTCACCACATGTTTCATGGCCTT 1379
85 eLeuValThrLysGluPheAsnSerIleMetGluIleLeuArgProTyrG 102
|||||
1380 TCTGGTGACCAAGAGTTTAACAGCATCATGAGATCTTCAGACCCCTAGC 1429
102 IValaPheTrpLeuThrAlaAlaPheCysIleLeuSerValLeuPheThr 118
|||||
1430 GCCCTTCTGGCTACCGCTCTCTGTATCCCTACGCGTCCCTTTTACG 1479
119 LeuThrPheValProGluThrLysGlyArgThrLeuGluGlnIlePheAl 135
|||||
1480 CTCACCTTTGTCCCTGAGACTAAAGGACGACTCTGGAACAATCACAGC 1529
135 aHisLeuArgAspGlyAspGlyProLeuSerValThrGlySerProGlu 151
|||||
1530 CCAATTGAGAGCGGTGACGACCCCTT TCTGTGACTGGCGACGCCCTGAG 1578
152 LeuSerTrpLeuArgValSerLysGlyValGluTrpProGln 165
|||||
1579 CTGAGGCTGTGGCTTCCGAGTTTCAAAAAGAGGTGAGTGGCCTCAG 1620

seq_name: gp_pat:AX076671

seq_documentation_block:
LOCUS AX076671 2072 bp DNA PAT 06-FEB-2001
DEFINITION Sequence 5 from Patent WO0104145.
ACCESSION AX076671
VERSION AX076671.1 GI:12711202

KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE 1 (bases 1 to 2072)
AUTHORS Thorens,B., Ibberson,M. and Uldry,M.
TITLE Glutx polypeptide family and nucleic acids encoding same
JOURNAL Patent: WO 0104145-A 5 18-JAN-2001;
University of Lausanne (CH)
FEATURES
source location/Qualifiers
1..2072
/organism="Mus sp."
/db_xref="taxon:10095"
21..1454
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28497.1"
/db_xref="GI:12711203"
/translation="MSPEDPOETOPLLRPPPEARTPRGRRVFLASFAALGLPLNFGFAL
GYSSPAIPSLRRTAPPALRLGDPAASWFGAVYTLGAAGVILGGMLDRSGRLSL
LCTVPTGFAVITTAARDVMMILGGRLTLGACVAVSLVAPVTSFIAPVAVRGLGSC
VOLMVTGILLAVAGVLEMRRLAVLGCVPPTMLLMCMPTPRFLLOHOYOE
AMALRFLMGSEGEHPVGAEGHGFOLAMLRRCVHKRPLITIGLWFOQSGVNAI
MEVANSIFEEAKFKDSSLASVTGIIQVLFVAVALIMDRAGRLKLLALSGVIMVFS
MSAFGTYFKLTQSGPNSHVLVPIIAEPADVHLIMDLAVGSMCLFIAGFVGMGPI
PMLMSEIFPLHIKGVATGCVLTNNMFALVTKEFNSVMEMLRPYGAFLMTAAFCAL
SVLFTLTVPETKGRITLQVTAHFEGR"

BASE COUNT 351 a 673 c 584 g 464 t

ORIGIN

alignment_scores:
Quality: 662.00 Length: 146
Ratio: 4.868 Gaps: 0
Percent Similarity: 93.151 Percent Identity: 84.932

alignment_block:
US-09-516-493-10 x AX076671 ..

Align seg 1/1 to: AX076671 from: 1 to: 2072

19 LeuValProIleSerAlaGluProAlaAspValHisLeuGlyLeuAlaTr 35
|||||
1086 CTGGTGGCCATCTCCGAGAGCCGTGATGTTCCCAAGTGGAGCATGGCCTG 1135
35 pLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTr 52
|||||
1136 GCTGGCTGTAGGAGCAATGTCCTTCATTCAGTTCGCGTGGGCT 1185
52 rPGLyProIleProTrpLeuLeuMetSerGluIlePheProLeuHisIle 68
|||||
1186 GGGGACCCATCCCTGGCTGCTCATGTCAAGATCTTCCCTGTCATGTC 1235
69 LysGlyValAlaThrGlyValCysValLeuThrAsnTrpPheMetAlaPhe 85
|||||
1236 AAGGCTGTGGCTACCGGACTGTGTCTCACCACATGTTTCATGGCCTT 1285
85 eLeuValThrLysGluPheAsnSerIleMetGluIleLeuArgProTyrG 102
|||||
1286 TCTGTGAGCCAAAGAGTTTAACAGCATCATGAGATGCTTCAGACCCCTAGC 1335
102 IValaPheTrpLeuThrAlaAlaPheCysIleLeuSerValLeuPheThr 118
|||||
1336 GTGGCTTCTGGCTACCGCTCTCTGTGCTTCAGTGTCTATTCACA 1385
119 LeuThrPheValProGluThrLysGlyArgThrLeuGluGlnIlePheAl 135
|||||
1386 CTGACCGTTGTCCCTGAGACTAAAGGCGAGCTGTGGAACAAGTCACAGC 1435
135 aHisLeuArgAspGlyAspGlyProLeuSerValThrGlySerProGlu 152
|||||
1436 CCAATTGAGAGCGAGTACAGACCCCTTCTGTGCTGGGAGCCCGGAGC 1485

152 euserTtrpleuAryalserLysGlyValGluTtrPro 164
 1486 TGAGCTGCTTTGGCTTTCAGAAAGAGTGGAGTGGCT 1523
 seq_name: gb_ro:MMU245936

seq_documentation_block: 2072 bp mRNA ROD 18-FEB-2000
 LOCUS MMU245936
 DEFINITION Mus musculus mRNA for glucose transporter (GLUTX1 gene).
 ACCESSION AJ245936
 VERSION AJ245936.1 GI:70186606
 KEYWORDS glucose transporter; GLUTX1 gene.
 SOURCE house mouse.
 ORGANISM Mus musculus.

REFERENCE
 AUTHORS 1 (bases 1 to 2072)
 TITLE Ibberson, M., Uldry, M., and Thorens, B.
 JOURNAL J. Biol. Chem. 275 (7), 4607-4612 (2000)
 MEDLINE 2018191
 REFERENCE 2 (bases 1 to 2072)
 AUTHORS Ibberson, M.R.
 TITLE Direct Submission
 JOURNAL Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
 and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
 CH-1005, SWITZERLAND

FEATURES
 source location/Qualifiers
 1..2072
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 21..1454
 /gene="GLUTX1"
 21..1454
 /gene="GLUTX1"
 /product="glucose transporter"
 /protein_id="CAI75719.1"
 /db_xref="GI:7018607"

gene
 CDS
 21..1454

BASE COUNT 351 a 673 c 584 g 464 t
 ORIGIN
 /translation="MSPEPQETOPILRPPREARTPRGRVFLASFAALGPIPLSFAL
 GTSPPAIPSLRTPALRLGDNASWFAVVTLAGAGGILGGLIDRSGRKLSLL
 CTVPEVTGFAVITAARDVMMILGRLTGLAGVASLAVYISLTAVPAYVRGLGSC
 VOLAVVTGILAVAGVLEWMLAVIGCVPTLMLIMCYMTERPFLITQHOYOA
 MAALRPLMGSEGEPEPVGAEHOGFOLALLRPGIKPLITIGISLWFOOLSGVNAI
 MYANSIFPEAKKXDSLSASVGIIOVLEFTVAALIMRAGRLLIATSGAVIMVPM
 SAFGTVEKTKQSLPSNSSHVLGVPITAEVDVYGLAMLAAGSMCLFIAGFVAGMPI
 PWLMSEIFPLHVGATGICVLINWPMALVYKERNVMEMLRPYGATWLTAAFCAL
 SVLEFTLVVPEETKGRILEQVTAHFEGR"

alignment_scores:
 quality: 662.00 Length: 146
 Ratio: 4.868 Gaps: 0
 Percent Similarity: 93.151 Percent Identity: 84.932

alignment_block:
 US-09-516-493-10 x MMU245936 ..
 Align seg 1/1 to: MMU245936 from: 1 to: 2072

19 leuValProIleSerAlaGluProAlaSerValHisLeuGlyLeuAlaTr 35
 1086 CTGGTGGCCATCGCGGAGCCTGTGATGTCCAGTGGAGCTGGCCTG 1135
 35 pLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyT 52
 1136 GCTGGCTGTAGCAGCATGTGCTCTTCATTCGCTGGCGGGGCT 1185
 52 rGlyIleProIleProIleuLeuMetSerGluIlePheProLeuHisIle 68

1186 GGGAGCCCAATCCCTGGCTCCTCATGTACAGATCTCCCTGCATGCTC 1235
 69 LysGlyValAlaThrGlyValCysValLeuThrSerTrpPheMetAlaPh 85
 1286 TCTAGTACCAAGACGTTCAACACCGTCATGAGATGCTGACCCATG 1335
 1236 AAGGTGTGGCTACCGCATCTGTCTCTCACACACGTGTTCAATGCTT 1285
 85 eLeuValThrLysGluPheAsnSerIleMetGluIleLeuArgProTyrG 102
 102 lYalAlaPheTrpLeuThrAlaAlaPheCysIleLeuSerValLeuPheThr 118
 1336 GTGCTTCTGCTCAGCCGCTGCTTCTGCTCAGATGCTCTATTCACA 1385
 119 LeuThrPheValProGluThrLysGlyLysArgThrLeuGluGlnIleThrAl 135
 1386 CTGACCGTGTGCTCGACACTAAAGGACGACTCTGACACAGTCACAGC 1435
 135 aHisLeuArgAspLysArgLysProLeuSerValThrLysSerProGluL 152
 1436 CCATTTCAGAGGAGGATGACAGACCCCTTCTGTGCTGAGACCCGAGC 1485
 152 euserTtrpleuAryalserLysGlyValGluTtrPro 164
 1486 TGAGCTGCTTTGGCTTTCAGAAAGAGTGGAGTGGCT 1523
 seq_name: gb_ro:AF232061

seq_documentation_block: 1843 bp mRNA ROD 23-JUN-2000
 LOCUS AF232061
 DEFINITION Mus musculus glucose transporter GLUT8 mRNA, complete cds.
 ACCESSION AF232061
 VERSION AF232061.1 GI:8671757
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus.

REFERENCE
 AUTHORS 1 (bases 1 to 1843)
 TITLE Carayannopoulos, M.O., Chi, M.M., Cui, Y., Pingsterhaus, J.M.,
 McKnight, R.A., Mueckler, M., Devaskar, S.U. and Moley, K.H.
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (13), 7313-7318 (2000)
 MEDLINE 20319023
 REFERENCE 2 (bases 1 to 1843)
 AUTHORS Moley, K.H., Carayannopoulos, M.O. and Cui, Y.
 TITLE Direct Submission
 JOURNAL Submitted (07-FEB-2000) OB/Gyn, Washington University, 4566 Scott
 Ave, St. Louis, MO 63110, USA

FEATURES
 source location/Qualifiers
 1..1843
 /organism="Mus musculus"
 /strain="129"
 /db_xref="taxon:10090"
 1..48
 49..1482
 /product="glucose transporter GLUT8"
 /protein_id="AF232061.1"
 /db_xref="GI:8671758"

5' UTR
 CDS

19 leuValProIleSerAlaGluProAlaSerValHisLeuGlyLeuAlaTr 35
 1086 CTGGTGGCCATCGCGGAGCCTGTGATGTCCAGTGGAGCTGGCCTG 1135
 35 pLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyT 52
 1136 GCTGGCTGTAGCAGCATGTGCTCTTCATTCGCTGGCGGGGCT 1185
 3' UTR
 BASE COUNT 291 a 609 c 529 g 410 t 4 others

ORIGIN

alignment_scores: Quality: 657.00 Length: 146
 Ratio: 4.831 Gaps: 0
 Percent Similarity: 93.151 Percent Identity: 84.247

alignment_block:
 US-09-516-493-10 x AF232061 ..

Align seg 1/1 to: AF232061 from: 1 to: 1843

```

19  LeuValProIleSerAlaGluProAlaAspValHisLeuGlyLeuAlaTr 35
   |||||.....|.....|.....|.....|.....|.....|.....|
1114 CTGTCGCCCATCGGGCGGACCGCTGTGATGTCCAGGGGACCTGGCTG 1163
   |||||.....|.....|.....|.....|.....|.....|.....|
35  pLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyT 52
   |||||.....|.....|.....|.....|.....|.....|.....|
1164 GTGCGCTGTAGCGACATGTGCTCTTCATGTGCTGGCTTCCGGTGGCT 1213
   |||||.....|.....|.....|.....|.....|.....|.....|
52  rpgIyProIleProTrrPleuLeuMetSerGluIlePheProLeuHisIle 68
   |||||.....|.....|.....|.....|.....|.....|.....|
1214 GGGACCCATCCCTGCTCCTCATGTACAGACTTCCTCTGATGTGC 1263
   |||||.....|.....|.....|.....|.....|.....|.....|
69  LysGlyAlaAlaThrGlyValCysValLeuThrAsnTrpPheMetAlaPh 85
   |||||.....|.....|.....|.....|.....|.....|.....|
1264 AAGGCTGTGCTACCGGACATCTGTCTCCACCACTGCTTCATGGCTT 1313
   |||||.....|.....|.....|.....|.....|.....|.....|
85  eLeuValThrLysGluPheAsnSerIleMetGluIleLeuArgProTyrG 102
   |||||.....|.....|.....|.....|.....|.....|.....|
1314 TCTAGTACCAAGAGATTACACAGCTCATGTGATGCTCAGACCCCTAC 1363
   |||||.....|.....|.....|.....|.....|.....|.....|
102  lYAlaPheTrpLeuThrAlaAlaPheCysIleLeuSerValLeuPheThr 118
   |||||.....|.....|.....|.....|.....|.....|.....|
1364 GTGCTTCTGCTCACCCTGCTCTTCTGCTCTCAGTGTCTTATTCACA 1413
   |||||.....|.....|.....|.....|.....|.....|.....|
119  LeuThrPheValProGluThrLysGlyArgThrLeuGluGlnIleThrAl 135
   |||||.....|.....|.....|.....|.....|.....|.....|
1414 CTGACCGTGTGCCCTGAGACTAAAGGCGAGACTCGAACAAGTACAGC 1463
   |||||.....|.....|.....|.....|.....|.....|.....|
135  aHisLeuArgAspGlyAspGlyProLeuSerValThrGlySerProGluL 152
   |||||.....|.....|.....|.....|.....|.....|.....|
1464 CCATTTCGAGGAGATGACAGACCCCTTCTGCTGGGAGAGCCCGAGC 1513
   |||||.....|.....|.....|.....|.....|.....|.....|
152  euSerTrpLeuArgValSerLysGlyValGluTrpPro 164
   |||||.....|.....|.....|.....|.....|.....|.....|
1514 TGACCTGCTTTGGGTTTCAGAGAGAGTGAGTGCCT 1551
   |||||.....|.....|.....|.....|.....|.....|.....|
seq_name: gb_ro:MMU17802
seq_documentation_block:
LOCUS      MMU17802      1490 bp      mRNA      13-MAY-2000
DEFINITION Mus musculus mRNA for glucose transporter 8 (GLUT8 gene).
ACCESSION  Y17802.1 GI:7688219
VERSION    Y17802.1 GI:7688219
KEYWORDS   glucose transporter 8; GLUT8 gene.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Mus.
REFERENCE  1 (bases 1 to 1490)
            Doege,H., Schumann,A., Bahrenberg,G., Brauers,A. and Joost,H.G.
            GLUT8, a novel member of the sugar transport facilitator family
            with glucose transport activity
            J. Biol. Chem. 275 (21), 16275-16280 (2000)
MEDLINE    20283667
REFERENCE  2 (bases 1 to 1490)
            Joost,H.G.
            Direct Submission
            Submitted (16-JUL-1998) H. Joost, Institute of Pharmacology and
            Toxicology, Technical University Aachen, Wendlingweg 2, D-52057
            Aachen, FRG

```

FEATURES

source Location/Qualifiers
 1..1490
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /tissue_type="testis"
 26..1459
 /gene="GLUT8"
 26..1459
 /gene="GLUT8"
 /codon_start=1
 /product="glucose transporter 8"
 /protein_id="CA89815.1"
 /db_xref="GI:7688220"
 /translation="MSPEDPOETPLRPPARTPRGRVFLASFAALGCLSPFGL
 GYSSPAISLRTPPALRLGDNAASFVAVVILGAAGILGMLDRGRRLSL
 CTPFTGFAVITARDVMMLIGRLTLGACVASLAVPYISLAIYPAVRGLGSC
 VOLMYVTGILAAVAGNVLFWRLAVLGCYPPRLMLLMCYMPEFRPLITHOYOA
 MAALRLWMSBEGWDEEPVGAHQFQDLALRRPGYIKPLITISIMVFOIGVNAI
 MFVANSIFEEARKFSDSLASVTVGIIQVLPFAVALIMDAGRRLILALSGVIMVSM
 SARGYIEKLITQSLPSNSHVLVPIAAEPDVQVGLAMLAAGSMCLFIAGFAVGWPI
 PMLMSEIPLHLKGVATGICVLTNMFMAFLVTKFESSVEMLRPYGAFWLTAFCAL
 SVLFTLTVEPFGRLTEQVTAHREGR"

BASE COUNT 222 a 499 c 438 g 331 t

ORIGIN

alignment_scores: Quality: 608.00 Length: 133
 Ratio: 4.864 Gaps: 0
 Percent Similarity: 93.985 Percent Identity: 84.962

alignment_block:
 US-09-516-493-10 x MMU17802 ..

Align seg 1/1 to: MMU17802 from: 1 to: 1490

```

19  LeuValProIleSerAlaGluProAlaAspValHisLeuGlyLeuAlaTr 35
   |||||.....|.....|.....|.....|.....|.....|.....|
1091 CTGTCGCCCATCGGGCGGACCGCTGTGATGTCCAAAGTGGAGCTGGCTG 1140
   |||||.....|.....|.....|.....|.....|.....|.....|
35  pLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyT 52
   |||||.....|.....|.....|.....|.....|.....|.....|
1141 GCTGCTGTAGGACACATGTGCTCTTCTCATTTGCTGCTTGGGAGGCT 1190
   |||||.....|.....|.....|.....|.....|.....|.....|
52  rpgIyProIleProTrrPleuLeuMetSerGluIlePheProLeuHisIle 68
   |||||.....|.....|.....|.....|.....|.....|.....|
1191 GGGACCCATCCCTGCGCTCTCATGTGACAGATCTTCCCTTGATGTGC 1240
   |||||.....|.....|.....|.....|.....|.....|.....|
69  LysGlyValAlaThrGlyValCysValLeuThrAsnTrpPheMetAlaPh 85
   |||||.....|.....|.....|.....|.....|.....|.....|
1241 AAGGCTGTGCTACCGGACATCTGTGCTCACCACACTGGTTCATGGCTT 1290
   |||||.....|.....|.....|.....|.....|.....|.....|
85  eLeuValThrLysGluPheAsnSerIleMetGluIleLeuArgProTyrG 102
   |||||.....|.....|.....|.....|.....|.....|.....|
1291 TCTAGTACCAAGAGATTACAGCGTCATGGAGATGCTCAGACCCCTACG 1340
   |||||.....|.....|.....|.....|.....|.....|.....|
102  lYAlaPheTrpLeuThrAlaAlaPheCysIleLeuSerValLeuPheThr 118
   |||||.....|.....|.....|.....|.....|.....|.....|
1341 GTGCTTCTGCTCACCCTGCTCTTCTGCTCTCATGCTCTCTATTCACA 1390
   |||||.....|.....|.....|.....|.....|.....|.....|
119  LeuThrPheValProGluThrLysGlyArgThrLeuGluGlnIleThrAl 135
   |||||.....|.....|.....|.....|.....|.....|.....|
1391 CTGACCGTGTGCCCTGAGACTAAAGGCGAGACTCGAACAAGTACAGC 1440
   |||||.....|.....|.....|.....|.....|.....|.....|
135  aHisLeuArgAspGlyAspGlyProLeuSerValThrGlySerProGlu 151
   |||||.....|.....|.....|.....|.....|.....|.....|
1441 CCATTTCGAGGAGATGACAGACCCCTTCTGTGCTGGGAGCCCGAG 1489
   |||||.....|.....|.....|.....|.....|.....|.....|
seq_name: gb_pat:AX179740
seq_documentation_block:
LOCUS      AX179740      2080 bp      DNA      06-AUG-2001
DEFINITION Sequence 28 from Patent WO0146258.

```



```

|||||.....|
1348 TGTGGCTTGCTCGCTTCTGCATCTTCAGTGCCTTTTCACCTTTTCC 1397
121 PheValProCylThrLysGlyArgThrLeuGluGlnIleThrAlaHisLe 137
1398 TGTGTCCCTCAAACTAAAGGAAGACTCTGGACAAATCACAGCCCATTT 1447
137 uArgAspGlyAspGlyProLeuSer 145
1448 TGAAGGGCGCATGACAGCCACTCACT 1472

seq_name: gb_om:AF321324

seq_documentation_block:
LOCUS AF321324 1012 bp mRNA MAM 02-JUL-2001
DEFINITION Bos taurus glucose transporter 8 (Glut8) mRNA, partial cds.
ACCESSION AF321324
VERSION AF321324.1 GI:14582715
KEYWORDS
SOURCE
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 1012)
Augustin,R., Navarrete-Santos,A. and Fischer,B.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2000) Anatomy and Cell Biology, Faculty of
Medicine, Martin Luther University Halle, Grosse Steinstrasse 52,
Halle 061097, Germany
FEATURES
source
location/Qualifiers
1..1012
/organism="Bos taurus"
/db_xref="taxon:9913"
<1..1012
/gene="Glut8"
<1..1005
/gene="Glut8"
/codon_start=1
/product="glucose transporter 8"
/protein_id="AAK69606.1"
/db_xref="GI:14582716"
/translation="1SETAYPEVRRLSCVOLNMTVLTLLAYLAWGLEWRLAVLG
CVAPSPMLLMCFMPEPRFLPLSHKHODAMAMQFLMWAOQWEPPLGAOHODPHV
AQLRRPGVYKPFILIGISLMAFQSLGVNVMVAETIFEEAFKDSLAVYGVGV
LFTAAALIMDRAGRRLTLTLLSGVNVSTSAFTGYFKLTGGSPSSSHVDLPALVSM
EAADTNVGLAMLVGNMCLFTAGFAVGNPFWLLMSLFLPLHYKCVATGCVLTNMF
MAFLVTKERSSLMELVLRPYGAFWLASAFICIFGVFLTLACVPETKGTLEDTAHEGR
"
BASE COUNT 153 a 323 c 302 g 234 t
ORIGIN
alignment_scores:
Quality: 546.00 Length: 123
Ratio: 4.667 Gaps: 1
Percent Similarity: 95.122 Percent Identity: 82.114

alignment_block:
US-09-516-493-10 x AF321324 ..
Align seg 1/1 to: AF321324 from: 1 to: 1012
22 ILSerAlaGluProAlaAspValHisLeuGlyLeuAlaTrpLeuAlaVal 38
:::| | | | | | | | | | | | | | | | | | | | | | | | | | |
646 GTCTCCATGAGAGCGCGCATACCAACGTTGGGCTGGCTGGCGGT 695
38 IGLySerMetCysLeuPheIleAlaGlyPheAlaValAlaGlyTrpGlyProI 55
| | | | | | | | | | | | | | | | | | | | | | | | | | |
696 GGGCAACATGTCCTTTCATCGCCGCTTGGCTGGCGGAGGCCCA 745
55 ILeProTrpLeuLeuMetSerGluIlePheProLeuHisIleLysGlyVal 71
| | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

746 TCCCGTGAGCTCCATGTCAGAGATCTTCCTTCGCATGTCAAGGGCGCTG 795
72 AlaThrGlyValCysValLeuThrAsnTrpPheMetAlaPheLeuValTh 88
796 GCCACGGCGCTCTGCGTCTCCACACACGTGTTTCATGGCTTTCTGGTAC 845
88 rLysGluPheAsnSerIleMetGluIleLeuArgProTrpGlyAlaPheT 105
846 CAAGAGTTTCAGACAGTCTCATGAGAGTGCTCAGGCGCTTACGCTTCT 895
105 rPheThrAlaAlaPheCysIleLeuSerValLeuPheThrLeuThrPhe 121
896 GGCTGGCTTCCTGCTTCGATCTTCGCTGCTGCTTTTACCTGGCGCTG 945
122 ValProGluThrLysGlyArgThrLeuGluGlnIleThrAlaHisLeuA 138
946 GTCCCTGAGACCAAAAGGAAGACTCTGGACAAATCACAGCCCATTTGA 995
138 rGAspGlyAspGlyPro 143
996 GGCCCGATGACAGCCCT 1012

seq_name: gb_pr:HS245937

seq_documentation_block:
LOCUS HS245937 1873 bp mRNA PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA for glucose transporter (GLUTX1) gene.
ACCESSION AJ245937
VERSION AJ245937.1 GI:7018305
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1873)
Ibberson,M., Uldry,M. and Thorens,B.
AUTHORS
TITLE GLUTX1, a novel mammalian glucose transporter expressed in the
central nervous system and insulin-sensitive tissues
JOURNAL J. Biol. Chem. 275 (7), 4607-4612 (2000)
2018191
MEDLINE
REFERENCE 2 (bases 1 to 1873)
Ibberson,M.R.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
CH-1005, SWITZERLAND
FEATURES
source
location/Qualifiers
1..1873
/organism="Homo sapiens"
/db_xref="taxon:9606"
4..1437
/gene="GLUTX1"
4..1437
/gene="GLUTX1"
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CAB75702.1"
/db_xref="GI:7018306"
/translation="MTPEDEPEERQPLIGPPGSAAPRGVRVFLAFAAALCPISFPA
LGTSSPAIPSLQRAAPAPRIDDAAASWFGAVVLTGAAGGVLGGLVDRAGRLSLIL
CSVPVAGFAVITTAADQVWMLLGGRLTGLAGCVAISLVAPLYSEIAPAVRGLLGGC
VOLMVVVGILLAVIAGVLEWRLAVLGVGPSIMLLMCFMPEPRFLTLTGRRDEA
MAALRFMGSEOGMEDPPTGAEOSEFLALIRPGIYKPFILGVSLMAFOOLSGVNAV
FAETTFEFAKFKEDSLASVYGVGVLTAAVALIMDRAGRRLTLTLLSGVNVSTSA
FTGYFKLTGGSPSSSHVDLPALVSM EAADTNVGLAMLVGNMCLFTAGFAVGNP
FWLLMSLFLPLHYKCVATGCVLTNMF MAFLVTKERSSLMELVLRPYGAFWLASAFIC
IFGVFLTLACVPETKGTLEDTAHEGR"
SVLFLFCVPELTKGRTLEDTAHEGR"
BASE COUNT 279 a 621 c 573 g 400 t
ORIGIN
alignment_scores:
Quality: 542.00 Length: 124

```

Ratio: 4.672 Caps: 0
Percent Similarity: 93.548 Percent Identity: 79.032
alignment_block:
US-09-516-493-10 x HSA245937 ..

Align seg 1/1 to: HSA245937 from: 1 to: 1873

```

21  ProlSerAlaGluProAlaSpValHisLeuGlyLeuAlaTrpLeuAl 37
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1075 CCTGCTCTGCACAGCTGTTGATGCCAGCGTGGGCTGGCTGGCTGGC 1124

37  aValGlySerMetCysLeuPheIleAlaGlyPheAlaValaGlyTrpGlyP 54
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1125 CGTGGGCAACATGTGCTCTTCATCGCCGCTTTCGGGTGGGCGGCGC 1174

54  rolleProTrpLeuMetSerGluIlePheProLeuHisIleGlyGly 70
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1175 CCATCCCTGGCTCTCATGTGACAGATCTTCCTCTGATGATCAAGGCGC 1224

71  ValAlaThrGlyValaCysValLeuThrAsnTrpPheMetAlaPheLeuVa 87
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1225 GTGGGCAAGGATGTGGCTCTCTCACCAAGCTGCTCATGGCTTCTGCT 1274

87  IThrGlyGluPheAsnSerIleMetGluIleLeuArgProGlyAlaP 104
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1275 GACCAAGAGATTCACAGACCTCATGAGGTCTCTGAGCCCTATGAGGCT 1324

104  heTrpLeuThrAlaAlaPheCysIleLeuSerValLeuPheThrLeuThr 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1325 TCTGGCTTGCTCCGCTTTCGATCTTCAGTGTCTTTCACCTTGTCTG 1374

121  PheValProGluThrGlyGlyArgThrLeuGluGluIleThrAlaHisIe 137
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1375 TGTGTCCCTGAATTAAGAAAGACCTCTGAGAACATCAGACCCATT 1424

137  uArgAspGlyAspGlyProLeu 144
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1425 TGAGGGCGCATGACAGCCACTC 1446

seq_name: gb_pat:AX076667

```

seq_documentation_block:
LOCUS AX076667 2217 bp DNA PAT 06-FEB-2001
DEFINITION Sequence 1 from Patent WO0104145.
ACCESSION AX076667
VERSION AX076667.1 GI:12711198
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2217)
AUTHORS Thorens,B., Ibberson,M. and Udry,M.
TITLE Glutk polypeptide family and nucleic acids encoding same
JOURNAL Patent: WO 0104145-A 1 18-JAN-2001;
University of Lausanne (CH)
FEATURES
source
1..2217
location/Qualifiers
CDS
/organism="Homo sapiens"
/db_xref="taxon:9606"
348..1781
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28495.1"
/db_xref="GI:12711198"
/translation="MTPEDEPETOPLPGPGSAPRCGRVETAAFAAALGLSFGFAL
GSSSPAIPLSQRAAPPAPRLDDAASMFAGAVVILGAAGVLGSMIVDRGRKLSLL
CSYFVAGFAVITAQDVWMLLGRLTIGACVASIVAVYSEIAYPAVRGLISGL
VOLMVVGLILAYLAGVLEWMLAVLVCVPSLMLLMCFMPEETPELILTOHRROA
MALRLMGSEOGMDPITGACDSFHILALROGCTKPTITGSLNAFQOLSVNAYM
FYAETIPEEKFRDSSIASVYVIVLFAVALIMDRGRGILVLVLSGVAVFST
ATGAYFKLTGGGPNSSHVAISAPVSAQPDVASGLAMLVGNMCLFIAGFAVGWPI

PWLMSEIFPLHVKGATGICVLTNMIMAPLVTKESSLMEVLRPYGAFWLASFCIF
SVLTPECVPEIKGKILEOITAEHGR
BASE COUNT 320 a 750 c 672 g 475 t
ORIGIN

alignment_scores:
Quality: 542.00 Length: 124
Ratio: 4.672 Caps: 0
Percent Similarity: 93.548 Percent Identity: 79.032

alignment_block:
US-09-516-493-10 x AX076667 ..

Align seg 1/1 to: AX076667 from: 1 to: 2217

```

21  ProlSerAlaGluProAlaSpValHisLeuGlyLeuAlaTrpLeuAl 37
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1419 CCTGCTCTGCACAGCTGTTGATGCCAGCGTGGGCTGGCTGGCTGGC 1468

37  aValGlySerMetCysLeuPheIleAlaGlyPheAlaValaGlyTrpGlyP 54
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1469 CGTGGGCAACATGTGCTCTTCATCGCCGCTTTCGGGTGGGCGGCGC 1518

54  rolleProTrpLeuMetSerGluIlePheProLeuHisIleGlyGly 70
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1519 CCATCCCTGGCTCTCATGTGACAGATCTTCCTCTGATGATCAAGGCGC 1568

71  ValAlaThrGlyValaCysValLeuThrAsnTrpPheMetAlaPheLeuVa 87
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1569 GTGGGCAAGGATGTGGCTCTCTCACCAAGCTGCTCATGGCTTCTGCT 1618

87  IThrGlyGluPheAsnSerIleMetGluIleLeuArgProGlyAlaP 104
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1619 GACCAAGAGATTCACAGACCTCATGAGGTCTCTGAGCCCTATGAGGCT 1668

1619  heTrpLeuThrAlaAlaPheCysIleLeuSerValLeuPheThrLeuThr 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1669 TCTGGCTTGCTCCGCTTTCGATCTTCAGTGTCTTTCACCTTGTCTG 1718

121  PheValProGluThrGlyGlyArgThrLeuGluGluIleThrAlaHisIe 137
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1719 TGTGTCCCTGAATTAAGAAAGACCTCTGAGAACATCAGACCCATT 1768

137  uArgAspGlyAspGlyProLeu 144
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1769 TGAGGGCGCATGACAGCCACTC 1790

seq_name: gb_pr: HSA011372

```

seq_documentation_block:
LOCUS HSA011372 2487 bp mRNA PRI 18-AUG-2000
DEFINITION Homo sapiens mRNA for sugar transporter (SLC2A6 gene).
ACCESSION AJ011372
VERSION AJ011372.2 GI:9843742
KEYWORDS SLC2A6 gene; sugar transporter.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2487)
AUTHORS Young,J.M., Woodward,K.J., Aziz,S., Burley,M., Kwiatkowski,D.J. and Povey,S.
TITLE Cloning of a sugar transporter gene, a G-beta subunit like gene and three novel genes in human chromosome 9q34
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2487)
AUTHORS Young,J.M.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-1998) Young J.M., MRC Human Biochemical Genetics Unit, University College London, Wolfson House, 4 Stephenson Way, London, NW1 2HE; UK
REMARK Revised by [3]


```

35 pLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValAlaGlyT 52
|||||
1148 GCTGGCTGTAGCAGCATGTGCTCTTCATCGAGTGGTTTTCAGACGAGCT 1197
|||||
52 rPgIlyProIleProTrrPleuLeuMetSerGluIlePheProLeuHisIle 68
|||||
1198 GGGGACCATCCCTGGCTCCTCATGTCAGAGATCTTCCCTCTGACATC 1247
|||||
69 LysGlyValAlaThrGlyValCysValLeuThrAsnTrpPheMetAlaPh 85
|||||
1248 AAGGCTGTGGTACCGGCTGTGTCTCTCACAACACTGTTCAATGCGCTT 1297
|||||
85 eLeuValThrLysGluPheAsnSerIleMetGluIleLeuArgProTyrG 102
|||||
1298 TCTGTGACCAAGAGTTTAAACATCATGAGATCCTCAGACCTCAG 1347
|||||
102 lYAlaPheTrpLeuThrAlaAlaPheCysIleLeuSerValLeuPheThr 118
|||||
1348 GCGGCTTCTGGCTCAGCGCTGCTGTATCTCAGCGCTCTTTTCAGC 1397
|||||
119 LeuThrPheValProGluThrLysGlyArgThrLeuGluIleThrAl 135
|||||
1398 CTCACCTTTGCTCCTGAGACTAAAGCGAGACTCTGGACAAATCAGACG 1447
|||||
135 aHisLeu.ArgAspGlyAspGlyProLeuSerValThrGlySerProGlu 151
|||||
1448 CCATTTGAGGAGCGGTGACGACCTT.TCTGTGACTGCGACGCTGAG 1496
|||||
152 LeuSerTrpLeuArgValSerLysGlyValGluTrpProGln 165
|||||
1497 CTGAGCTGGCTTGGGTTTCAAAAGAGTGAGTGGCTCAG 1538
|||||
seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF55867

```

```

seq_documentation_block:
ID AAF55867 standard; cDNA: 2072 BP.
XX
AC AAF55867;
XX
DT 17-APR-2001 (first entry)
XX
DE Murine GLUTX1 coding sequence.
XX
KW Murine; GLUTX; gene therapy; vaccine; hexose transport modulator;
KW hexose transport disorder; ischemia; diabetes; hyperglycaemia; ss;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS Mus sp.
XX
PN WO200104145-A2.
XX
PD 18-JAN-2001.
XX
PF 14-JUL-2000; 2000MO-IB01042.
XX
PR 14-JUL-1999; 9905-0143907.
PR 27-AUG-1999; 9905-0151140.
PR 23-FEB-2000; 2000US-0184285.
PR 13-JUL-2000; 2000US-0616132.
XX
PA (UYLA-) UNIV LAUSANNE.
XX
PI Thorens B, Ibberson M, Uldry M;
XX
DR WPI: 2001-112615/12.
DR P-PSDB: AAB66934.
XX
XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
PT e.g. ischemia and diabetes -
XX
PS Claim 3; page 73-74; 124pp; English.
XX

```

```

CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is the coding sequence
CC for murine GLUTX1.
XX
SQ Sequence 2072 BP; 351 A; 673 C; 584 G; 464 T; 0 other;

```

```

alignment_scores:
Quality: 662.00 Length: 146
Ratio: 4.868 Gaps: 0
Percent Similarity: 93.151 Percent Identity: 84.932

```

```

alignment_block:
US-09-516-493-10 x AAF55867 ..

```

```

Align seg 1/1 to: AAF55867 from: 1 to: 2072

```

```

19 LeuValProIleSerAlaGluProAlaAspValHisLeuGlyLeuAlaTr 35
|||||
1086 CTGGTGGCCATCGCGGCGAGCTGTGTGATGTCAGAGTGGAGCTGGCTG 1135
|||||
35 pLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValAlaGlyT 52
|||||
1136 GCTGGCTGTAGCAGCATGTGCTCTTCATGCTGCTTGGCGTGGGCT 1185
|||||
52 rPgIlyProIleProTrrPleuLeuMetSerGluIlePheProLeuHisIle 68
|||||
1186 GGGGACCATCCCTGGCTCCTCATGTCAGAGATCTTCCCTCGCATGTC 1235
|||||
69 LysGlyValAlaThrGlyValCysValLeuThrAsnTrpPheMetAlaPh 85
|||||
1236 AAGGTTGTGGCTACCGGATCTGTGTCTCACCACACTGTTGATGAGCTT 1285
|||||
85 eLeuValThrLysGluPheAsnSerIleMetGluIleLeuArgProTyrG 102
|||||
1286 TCTAGTGAACAAAGAGTTCAACAGCGCTCATGAGATGCTCAGACCTCAG 1335
|||||
102 lYAlaPheTrpLeuThrAlaAlaPheCysIleLeuSerValLeuPheThr 118
|||||
1336 GTGCTTGTGGCTCAGCGCTGCTTGTGCGCTCAGTGTCTTATTCACA 1385
|||||
119 LeuThrPheValProGluThrLysGlyArgThrLeuGluIleThrAl 135
|||||
1386 CTGACCGTTGTCTGAGACTAAAGCAGAGCTGTGACAGATCAGACG 1435
|||||
135 aHisLeuArgAspGlyAspGlyProLeuSerValThrGlySerProGluL 152
|||||
1436 CCATTTGAGGAGCATGACAGACCTTCTGTGCTGGAGACCCCGAGC 1485
|||||
152 eUSeTrTrpLeuArgValSerLysGlyValGluTrpPro 164
|||||
1486 TGAGCTGCTTGGGTTTCAGAACGAGTGAGTGGCT 1523
|||||
seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAD09552

```

```

seq_documentation_block:
ID AAD09552 standard; cDNA: 2080 BP.
XX
AC AAD09552;
XX
DT 10-SEP-2001 (first entry)
XX
XX Human transporter and ion channel-1 (TRICH-1) cDNA.
DE Human; transporter and ion channel-1; TRICH-1; vaccine; cystic fibrosis;
KW gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy;
KW hypertension; angina; neurological disorder; asthma; bipolar disorder;
KW dementia; depression; Alzheimer's disease; epilepsy; mood; arrhythmia;
KW

```

XX	Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke;
KW	Huntington's disease; Parkinson's disease; cerebellar neoplasm; allergy;
KW	demyelinating disease; mental disorder; Schizophrenia; polymyositis;
KW	muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease;
KW	dermatomyositis; diabetes mellitus; immunological disorder; psoriasis;
KW	rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus;
KW	sickle cell anaemia; Wilson's disease; infertility; Cushing's disease;
KW	scleroderma; pulmonary artery stenosis; noctropic; Addison's disease;
KW	malabsorption syndrome; hypercholesterolaemia; cancer; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	46..1479
FT	/*tag=a
FT	/product="Human TRICH-1 protein"
XX	
PN	W0200146258-A2.
PD	28-JUN-2001.
XX	
PF	22-DEC-2000; 2000WO-US35095.
XX	
PR	23-DEC-1999; 99US-0172000.
PR	14-JAN-2000; 2000US-0176083.
PR	21-JAN-2000; 2000US-0177332.
PR	28-JAN-2000; 2000US-0178572.
PR	02-FEB-2000; 2000US-0179758.
PR	10-FEB-2000; 2000US-0181625.
XX	
PA	(INCYT) INCYTE GENOMICS INC.
XX	
PI	Baughn MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy R, Lal P;
PI	Hillman Jr, Azimzai Y, Yue H, Nguyen DB, Yao MG, Gandhi AR;
PI	Tang YT, Khan FX.
XX	
DR	WPI: 2001-418042/4.
DR	P-PSDB: AA004888.
XX	
PT	Novel human transporter and ion channel proteins useful for treating
PT	and preventing transport, neurological, muscle and immunological
PT	disorders -
XX	
XX	Claim 5; Page 140-141; 160pp; English.
XX	
PS	The present sequence is transporter and ion channel-1 (TRICH-1) cDNA.
CC	TRICH is used as vaccine. TRICH is useful for treating a disease or
CC	condition associated with decreased expression of functional TRICH,
CC	such as transport disorder including amyotrophic lateral sclerosis,
CC	cystic fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth
CC	disease, Duchenne muscular dystrophy, angina and hypertension,
CC	neurological disorders including Alzheimer's disease, amnesia, bipolar
CC	disorder, dementia, depression, epilepsy, ischaemic cerebrovascular
CC	disease, stroke, cerebral neoplasms, Pick's disease, Huntington's
CC	disease and Parkinson's disease, demyelinating diseases, mental disorders
CC	including mood, anxiety, Schizophrenia and seasonal affective disorder,
CC	muscle disorder including cardiomyopathy, myocarditis, polymyositis,
CC	dermatomyositis, arrhythmias and asthma and immunological disorders
CC	including AIDS, adult respiratory distress syndrome (ARDS), allergies,
CC	anemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
CC	syndrome, systemic lupus erythematosus and other diseases including
CC	sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary
CC	artery stenosis, Grave's disease, Cushing's disease, Addison's disease,
CC	glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers
CC	psoriasis and viral, bacterial, fungal, helminthic and protozoal
CC	infections. TRICH DNA is useful in gene therapy and in diagnostic
CC	purposes.
XX	
XX	Sequence 2080 BP; 300 A; 668 C; 632 G; 460 T; 0 other;

```
alignment_scores:
  Quality: 552.00      Length: 125
```

Ratio:	4.678	Gaps:	0
Percent Similarity:	94.400	Percent Identity:	80.000
alignment_block:			

Align seg 1/1 to: AAD09552 from: 1 to: 2080

21 ProIleSerAlaGluProAlaAspValHisLeuGlyLeuAlaTrpLeuAl 37
1117 CSTGTCTCTGCACAGCCTGTTCATGCCAGCGTGGGGCTGGCCCTGGCTGGC 1166

37 aValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpCylP 54
1167 CGTGGCAGCATGTGCCTCTTCAATCCGCCGCTTTGCGGTGGGCTGGGGC 1216

54 rolleProTrpleuLeuMetSerGluIlePheProLeuHisIleLysGly 70
|||||
1217 CCATCCCTGGCTCCTCATGTGCAGATCTTCCCTGTGCATGTCAAGGCG 1266

71 ValAlaThrGlyValCysValLeuThrAsnTrpMetAlaPheLeuVal 87
 |||||
 1267 GTCGCACAGGCATCTGCGTCTCTACCACTGGCTCATGGCCTTTCCTGT 1316

87 IThlySGluPheAsnSerIleMetGluIleLeuArgProTyrGlyAlaP 104
 |||||
 |||||
 1317 GACCAAGGAGTTTCAGCAGCCTCATGTGAGGTCTCTCAGGCCCTATGAGCCT 1366

```

104  he|tr|pleu|thr|Ala|Ala|Phe|cys|Ile|eu|Ser|Val|Leu|Phe|thr|Leu|thr  120
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1367  TC|GGC|TTG|CCT|CCG|C|TTT|CT|GCA|CTT|CA|G|T|G|C|T|TT|CA|C|TTT|G|TT|C  1416

```

121 phevalproglutThrLysGlyArgThrLeuGluGlnIleThrAlaHisLe 137
 |||||
 1417 TCGNGTCCCTGGAACATAAAGGAAGACTCTGGAAACAATCACAGCCCATTT 1466

```

137 uargaspGlyaspGlyProLeuSer 145
      :::: |||||:::
1467 TGAGGGCGCATGACAGCCACTCACT 1491

```

```
seq_name: /SIDS2/gcgcdata/geneseq/geneseq/NA2001.DAT:AAf55865
seq_documentation_block:
  10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 10
```

XX
AC
XX
AAAF55865;
17 000 0001

XX	Human GLUTX1 coding sequence
DE	
XX	

KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease
XX

XX
PN
XX
DD

WO200104145-A2.
18-JAN-2001

XX	14-JUL-2000; 2000WO-1B01042.
PF	
XX	
14 JUL 2000	0000 0142007

PR	27-AUG-1999;	99US-0151140.
PR	23-FEB-2000;	2000US-0184285.
PR	13-JUL-2000;	2000US-0616132.

AA
PA (UYLA-) UNIV LAUSANNE

PI Thorens B, Ibberson M, Uldry M;
XX
DR WP1; 2001-112615/12.


```

18 uLeuValProIleSerAlaGluProAlaAspValHisLeuGly...LeuA 34
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
215 GTCTGGGGGAGCTTGGCCGAGCGACGACCGCTGACCTCA 264
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
34 IaTrpLeuAlaValAlaGlySerMetCysLeuPheIleAlaGlyPheAlaVal 50
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
265 CCGTGGGTCCTGCTGGCCACCATGCTTTCATCATGATGAGGCTACGCCGCTG 314
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
51 G1YTrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLeuHis 67
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
315 GGCTGGGGTCCCATCACCCTGGCTCATGCTGAGGCTCGGCCCGCTGCG 364
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
67 sIleLysGlyValAlaThrGlyValCysValLeuThrAsnTrpPheMetA 84
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
365 TGCCCGGGGCTGGCCCTGACGGGCTGCTGCGCAGCGCTGCTCAACG 414
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
84 IaPheLeuValThrLysGluPheAsnSerIleMetGluIleLeuValArgPro 100
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
415 CCTTCGTCCTACCAAGTCTCTGCTGCCAGTGGT.GAGCACCTTGGGCT 463
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
101 TyrGlyAlaPheTrpLeuThrAlaAlaPheCysIleLeuSerValLeuP 117
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
464 CCAAGTCCCTTCTCTCTTCTTCCGGCCATCTGCTGTGAGCTGGCT 513
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
117 heThrLeuThrPheValProGluThrLysGlyArgThrLeuGluGlnIle 133
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
514 TCACAGGCTGCTGTCGCCGACCAAGGACGCTCCCTGGACACAGATC 563
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
134 ThrAlaHisLeuValArgAsp.GlyAspGlyPro.LeuSerValThrGlySer 149
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
564 GAGTCCCTTCTTCCGACGAGGAGAGGTCCTTCTGCGCTAGGTCACAGCT 613
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
150 ProGluLeuSer.....TrrPle 155
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
614 CCCCCTGAGAGGGGCCAACCACCGAGTGGCTGCTGTGTGGCT 663
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
155 uArg 156
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
664 ACAA 667

seq_name: /sids2/gcgdata/geneseq/geneseqn/NA2001.DAT.AAF55870
seq_documentation_block:
ID   AAF55870 standard; cDNA: 1541 BP.
AC   AAF55870;
XX
XX
XX   17-APR-2001 (first entry)
XX
XX
XX   Human GLUTX3 coding sequence.
DE
XX
XX   Human: GLUTX: gene therapy: vaccine: hexose transport modulator:
KM   hexose transport disorder: ischaemia; diabetes; hyperglycaemia; ss:
KM   hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
OS   Homo sapiens.
XX
XX   WO200104145-A2.
XX
XX   18-JAN-2001.
XX
XX
XX   14-JUL-2000; 2000WO-1B01042.
XX
XX
XX   14-JUL-1999; 99US-0143907.
XX   27-AUG-1999; 99US-0151140.
XX   23-FEB-2000; 2000US-0184285.
XX   13-JUL-2000; 2000US-0616132.
XX
XX   (UYLA-) UNIV LAUSANNE.
XX
XX   Thorens B, Idberson M, Uldry M;
XX
XX   WPI; 2001-112615/12.

```

```

DR   P-PSDB; AAB66937.
XX
XX   Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT   the prevention, diagnosis and treatment of hexose transport disorders,
PT   e.g. ischemia and diabetes -
XX
XX   Claim 3; Page 81-82; 124pp; English.
XX
XX   The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC   AAB66932-AAB6941). The GLUTX proteins are related to the facilitative
CC   glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC   function. The GLUTX proteins may be used in the diagnosis, prevention and
CC   treatment of hexose transport disorders such as ischaemia, diabetes,
CC   hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC   neurodegenerative disease. The present sequence is the coding sequence
CC   for human GLUTX3.
XX
XX   S0
XX   Sequence 1541 BP; 209 A; 536 C; 482 G; 311 T; 3 other;

alignment_scores:
Quality: 285.00      Length: 145
Ratio: 2.689        Gaps: 4
Percent Similarity: 73.103      Percent Identity: 48.966

alignment_block:
US-09-516-493-10 x AAF55870 ..

Align seg 1/1 to: AAF55870 from: 1 to: 1541

2 ArgProLeu**Asn**TrpIleProArgAlaAlaGlyIleArgHisG1 18
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1103 AGGCTCGAGCCCAAC.....AGCACTGGGGCTGGAAAGCGA 1143
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
18 uLeuValProIleSerAlaGluProAlaAspValHisLeuGly...LeuA 34
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1144 GTCCCTGGGGGACTTGGCGCAGCCCTGGCAGACACCCCTGACCTCA 1193
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
34 IaTrpLeuAlaValAlaGlySerMetCysLeuPheIleAlaGlyPheAlaVal 50
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1194 CCGTGGTCCCTGCTGGCCACCATGCTTTCATCATGAGGCTACGCCGCTG 1243
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1244 GGCTGGGGTCCCATCACCCTGCTGCTCATGCTGAGGCTGACGCCCTGG 1293
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
67 sIleLysGlyValAlaThrGlyValCysValLeuThrAsnTrpPheMetA 84
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1294 TGCCCGGGGCTGGCCCTGACGGGCTGCTGCGCAGCGCTGCTCAACG 1343
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
84 IaPheLeuValThrLysGluPheAsnSerIleMetGluIleLeuValArgPro 100
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1344 CTTTCGTCCTACCAAGTCTCTGCTGCCAGTGGT.GAGCACCTTGGGCT 1392
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
101 TyrGlyAlaPheTrpLeuThrAlaAlaPheCysIleLeuSerValLeuP 117
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1393 CCAGTGGCTTTCCTCTTCTGCGGCCATCTGCTGTGAGGCTGGCTG 1442
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
117 heThrLeuThrPheValProGluThrLysGlyArgThrLeuGluGlnIle 133
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1443 TCACAGGCTGCTGTCGCCGACCAAGGACGCTCCCTGGACACAGATC 1492
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
134 ThrAlaHisLeuValArgAsp.GlyAspGlyPro 143
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1493 GAGTCCCTTCTTCCGACGAGGAGAGGTCCT 1523
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

seq_name: /sids2/gcgdata/geneseq/geneseqn/NA2001.DAT.AAF67134
seq_documentation_block:
ID   AAF67134 standard; cDNA: 378 BP.
AC   AAF67134;
XX
XX
XX

```

DT 09-APR-2001 (first entry)
 XX Novel human polynucleotide, SEQ ID NO: 2890.
 DE
 XX
 KW Human: cytostatic; gene therapy: colon cancer; prostate cancer;
 KW Breast cancer; Lung cancer; cancer detection; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200102568-A2.
 XX
 PD 11-JAN-2001.
 XX
 PF 30-JUN-2000; 2000WO-US18374.
 XX
 PR 02-JUL-1999; 99US-0142310.
 PR 02-JUL-1999; 99US-0142311.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A,
 PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
 PI Ckenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
 PI Kita D, Garcia V, Jones LW, Strache-Grain B;
 XX
 DR WPI: 2001-091805/10.
 XX
 PT Library of polynucleotides for diagnosing a cancerous state of a
 PT mammalian cell and detecting cancer, particularly of the colon or
 PT prostate, comprises 3351 human polynucleotide sequences .
 XX
 PS Claim 9: Page 974; 1046pp; English.

The present sequence is one of 3351 sequences in a library of human
 CC polynucleotides. The library is used to detect differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell and can
 CC detect colon, prostate, breast and lung cancer. The library can be used
 CC to produce probes for detection of mRNA and to produce additional copies
 CC of the polynucleotides. The probes can be used for chromosome mapping of
 CC the polynucleotide and for detection of transcription levels. Ribozymes
 CC or antisense oligonucleotides can be generated. The polynucleotides and
 CC their gene products are used as genetic or biochemical markers (e.g. in
 CC blood or tissues) that will detect the earliest changes along the
 CC carcinogenesis pathway and/or monitor the efficacy of therapies and
 CC preventive interventions. The polynucleotides, polypeptides and
 CC antibodies against them can be used in pharmaceutical compositions to
 CC treat the cancers and proliferative disorders such as neoplasia,
 CC dysplasia and hyperplasia.
 XX
 SO Sequence 378 BP; 45 A; 130 C; 117 G; 86 T; 0 other;

alignment_scores:
 Quality: 281.00 Length: 103
 Ratio: 3.306 Gaps: 2
 Percent Similarity: 82.524 Percent Identity: 59.223

alignment_block:
 US-09-516-493-10 x AAF67134 ..

Align seg 1/1 to: AAF67134 from: 1 to: 378

43 LeuHeiLeaLaglyPheAlaValGlyTrpGlyProIleProIlePhe 59
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 31 CTCCTTCATCATGGGCTACGCCGTGGGCTGCCATCACCCTGGCTCT 80
 59 uMeSerGluIlePheProLeuHisIleLysGlyValAlaThrGlyValC 76
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 81 CATCTCGAGGTCCTGCCCTGGCGTGGCGCTGGCGCTCAGAGGCTCT 130
 76 yValLeuThrAsnTrpMetAlaPheLeuValThrlYsGluPheAsn 92
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

131 GCGTGTGSCAGCTGGCTACCGGCTGCTCCTCACCAAGTCTCTCTG 180
 93 SerIleMeGluIleLeuArgProTrpGlyAlaPheTrpLeuThr AlaA 109
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 181 CCAAGTGTGAGACACCTTCGCTCCAGGCTGCTCTCTCTCTCTCTGCGG 229
 109 lAPheCysIleLeuSerValLeuPheTrpLeuThrPheValProGluTrp 125
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 230 CCATCTGCTGTGTGAGCCCTGCTGTTCACAGGCTGCTCTCTCTCTCTGCGG 279
 126 LysGlyArgThrLeuGluGlnIleThrAlaHisLeuArgAsp. GlyAspG 142
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 280 AAGGACGGTCCCTGGAGCAGATCGATCTTCTTCCGACGCGGAGAAG 329
 142 lYPro 143
 |||||
 330 GTCTCT 334

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA23461

seq_documentation_block:

ID AAA23461 standard; cDNA; 823 BP.
 XX
 AC AAA23461:
 XX
 DT 19-JUN-2000 (first entry)
 XX
 DE cDNA encoding human secreted protein vq1_1, SFG ID NO:77.
 XX
 KW Human: secreted protein; cancer; tumour; cardiovascular disorder;
 KW blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;
 KW infection; fungal; bacterial; viral; HIV; allergy; arthritis;
 KW neurodegenerative disease; asthma; contraceptive; ss.
 XX
 OS Homo sapiens.
 XX

Key Location/Qualifiers
 FT 54..371
 FT CDS /*tag= a
 FT /product= "Human secreted protein vq1_1"
 XX
 PN W0200011015-A1.
 XX
 PD 02-MAR-2000.
 XX
 PF 24-AUG-1999; 99WO-US19351.
 XX
 PR 24-AUG-1998; 98US-0097638.
 PR 24-AUG-1998; 98US-0097659.
 PR 09-SEP-1998; 98US-0099618.
 PR 28-SEP-1998; 98US-0102092.
 PR 25-NOV-1998; 98US-0109978.
 PR 23-DEC-1998; 98US-0113645.
 PR 23-DEC-1998; 98US-0113646.
 PR 23-AUG-1999; 99US-0379246.
 XX
 PA (ALPH-) ALPHAGEN INC.
 XX
 PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
 DR WPI: 2000-224657/19.
 DR P-PSDB; AAY95019.
 XX
 PT New secreted or transmembrane proteins and polynucleotides encoding
 PT them, useful for treating neurodegenerative disorders, autoimmune
 PT diseases and cancer -
 XX
 PS Claim 86; Page 335; 357pp; English.
 XX

The invention relates to 40 human secreted proteins (AAY94981-Y95020),
 CC and cDNA sequences encoding them (AAA23423-A23462). The secreted
 CC proteins of the invention include those that are thought to be only
 CC partially secreted, i.e., transmembrane proteins. The proteins of the

[illegible]

PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139761.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 24-JUN-1999; 9905-0140354.
PR 28-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0143297.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 16-JUL-1999; 9905-0144087.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144684.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.

PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 13-OCT-1999; 9905-0159296.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.

[illegible]

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144662.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151086.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152353.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.


```

1800 ACCGACAGACTGATGATTAGGGCCAGATGAGATGACCCGGC 1849
156 .....ArgValSerLygLyValu 162
1850 CCCGGATCCACCGGTCGCCACCATGTGACGACGCGAGG 1893
seq_name: /cgn2_6/plodata/2/lna/5B_COMB.seq:US-08-928-692-9

```

seq_documentation block:

```

; Sequence 9, Application US/08928692
; Patent No. 5958727
; GENERAL INFORMATION:
; APPLICANT: Brody, Howard
; APPLICANT: Yaver, Deborah S.
; APPLICANT: Lamsa, Michael
; APPLICANT: Hansen, Kim
; TITLE OF INVENTION: Methods for Modifying the Production of
; NUMBER OF INVENTION: a Polypeptide
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 59587270 No. 5958727disk of No. 5958727th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,692
; FILING DATE: 12-SEPT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4944,200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3000 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-928-692-9

```

alignment_scores:

```

Quality: 152.50 Length: 101
Ratio: 2.210 Gaps: 1
Percent Similarity: 68.317 Percent Identity: 33.663

```

alignment_block:

```

US-09-516-493-10 x US-08-928-692-9 ..

```

```

Align seg 1/1 to: US-08-928-692-9 from: 1 to: 3000

```

```

29 ValHisLeuGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPhe1 45
:.....:
1706 ATCTACTTCAACCTGAAGACCTTATCCGACATAGCAGTTCTTACTTTTCGT 1755
45 ealaglyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerg 62
:.....:
1756 TGCTTCTTCCGCGCGGTCTAGGCCAGTCCCTTCATTTTAGGCTCTG 1805
62 IuilePheProLeuHisIleLygLyValAlaThnGlyValAlcysValleu 78
:.....:

```

```

1806 AACTCGTTCGCCCGAGGCTGTCCGCCGACAGAGAGCTGGCGCTGGGA 1855
79 ThrAsnTrpPheMetAlaPheLeuValThrLygLeuPheAsnSerIle 95
:.....:
1856 GCGAAGCTGGATTGCGACGTTTCATCGTGACACATTTTTCGATGTTAA 1905
95 tGluLeuLeuArgProLygLy...AlaPheTrpLeuThrAlaAlaPheC 111
:.....:
1906 CGATTTGTTGGCGGCGGACGACGACATCTACTGATCTTTCGACGCGATG 1955
111 ystIleLeuSerValLeuPheThrLeuThrPheValProGluThrLygLy 127
:.....:
1956 CCGTCTCCGCGAGATTTCATCTACTGCTGTGGTGCCGACGACCAAGGG 2005
128 Arg 128
:
2006 AAG 2008

```

```

seq_name: /cgn2_6/plodata/2/lna/6B_COMB.seq:US-09-103-840A-1

```

seq_documentation block:

```

; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
; US-09-103-840A-1

```

alignment_scores:

```

Quality: 149.00 Length: 133
Ratio: 1.713 Gaps: 4
Percent Similarity: 65.414 Percent Identity: 29.323

```

alignment_block:

```

US-09-516-493-10 x US-09-103-840A-1 ..

```

```

Align seg 1/1 to: US-09-103-840A-1 from: 1 to: 4411529

```

```

19 LeuValProIleSerAlaGluProAlaAspValHisLeuGlyLeuAlaTr 35
:.....:
3718187CTGATCACCGTATTCGCCACACCTCCGATGCTGCGCGGCTGTG.. 3718234
35 pleuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyT 52
:.....:
3718235..TTGGGGTTCCGCCGCGTCTGCTGTCATTCATTCAGGATTCAGT 3718283
52 rglIleProIleProTrpLeuLeuMetSergIuilePheProLeuHisIle 68
:.....:
3718284TTCGGCTCGTGTCTGTGGGTGACGCGCGGACAGCTTCCGTCGCCGCTG 3718333
69 lvsGlyValAlaThnGlyValAlcysValLeuThrAsnTrpPheMetAlaPh 85
:.....:
3718334CGGTGATGATGATGAGGCCG...ATGCTCACCTCGACACACTGACGCCAA 3718380
85 eleuValThrLygLeuPheAsn...SerIleMetGluLeuLeuArgProT 101
:.....:
3718381CCGATCGTTCGCCGCTTCTCGTCCATCAGATCGCTGCTGTGCGCGCG 3718430

```

```

101 yrcgYlaIaPhETrPLeuThrAlaIaIaPhEcysIleLeuSerValLeuPhe 117
118 ThrLeuThrPheValProGluThrLysGluYArgThrLeuGluGlnIleTh 134
3718481GTCACCGCTTTCGGCAGTTCGGCGCTTCGCGCTTCGTCGTCGTCG 3718480
3718481GTCACCGCTTTCGGCAGTTCGGCGCTTCGCGCTTCGCGCTTCGTCG 3718528
134 rAlaHisLeuArGAspGlyAspGlyProLeuSerValThrGlySerPro 150
3718529.CGCGCACTTCGGAGAACGGCGCGCTTCGCGCGCGAGCGGTCACCG 3718576
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-998-416-719

seq_documentation_block:
: Sequence 719, Application US/08998416
: Patent No. 6239264
: GENERAL INFORMATION:
: APPLICANT: Philippson, Peter
: APPLICANT: Pohlmann, Rainer
: APPLICANT: Steiner, Sabine
: APPLICANT: Mohr, Christine
: APPLICANT: Wendland, Jurgen
: APPLICANT: Knechtle, Philipp
: APPLICANT: Redtschung, Corinne
: TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBVA GOSYP11
: NUMBER OF SEQUENCES: 1152
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 6239264artis Corporation
: STREET: 3054 Cornwallis Road
: CITY: Research Triangle Park
: STATE: No. 6239264th Carolina
: COUNTRY: USA
: ZIP: 27709
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/998,416
: FILING DATE: 24-DEC-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: CH 0016/97
: FILING DATE: 31-DEC-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Weigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-541-8587
: TELEFAX: 919-541-8689
: INFORMATION FOR SEQ ID NO: 719:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 708 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: PAG1471RP
US-08-998-416-719

alignment_scores:
: quality: 146.50 Length: 103
: Ratio: 2.063 Gaps: 2
: Percent Similarity: 68.932 Percent Identity: 33.010

alignment_block:
US-09-516-493-10 x US-08-998-416-719/rev ..

```

```

Align seq 1/1 to reverse of: US-08-998-416-719 from: 1 to: 708
42 CysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpL 58
536 TGCCCTACATTTTCTTCTTTCGACACTACCTGGGCCCAATTCGCTATG 487
58 euleMetSerGluIlePheProLeuHisIleLysGlyValAlaThrGly 74
486 TCATTGTTTCGAGACTTCCCATTAAGATCAAGCCAGGCGATGGCT 437
75 ValCysValLeuThrAsnTrpPheMetAlaPheLeuValThrLysGluPh 91
436 ATCGCCACAGCGCGGAATGTGATAGGGCTTTTGTGATCGCTCTTAC 387
91 eAsnSerIleMetGluIleLeuArGProTyGlyAlaIaPhETrPLeuThr 108
386 ACCATACATTTACCAATGGCATCAGGTTTACTACCGCTACGTTTCATGG 337
108 lAlaIaPhEcysIleLeuSerValLeuPheThrLeuThrPheValProGlu 124
336 GTGTTTG..ATCTTTCGTTCTTCTACATCTTCTTGTCCCGCAG 290
125 ThrLysGlyArGThrLeuGluGlnIleThrAlaHisLeuArGAspGlyAs 141
289 ACAAGCGTTTGACTTTCGAGAGGTGCGAGAGATGTCGACGAAAGTGT 240
141 pGlyPro 143
239 GGTACCT 233

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-031-392-1

seq_documentation_block:
: Sequence 1, Application US/09031392
: Patent No. 5942398
: GENERAL INFORMATION:
: APPLICANT: Tartaglia, Louis A.
: APPLICANT: Weng, Xun
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES
: TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM compatible
: OPERATING SYSTEM: Windows95
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/031,392
: FILING DATE: 26-FEB-1998
: ATTORNEY/AGENT INFORMATION:
: NAME: Melikiejohn, Ph.D., Anita L.
: REGISTRATION NUMBER: 35,283
: REFERENCE/DOCKET NUMBER: 07334/072001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617/542-5070
: TELEFAX: 617/542-8906
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2343 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:

```

```

; NAME/KEY: Coding Sequence
; LOCATION: 73...1761
US-09-031-392-1

```

Quality:	118.50	Length:	141
Ratio:	1.519	Gaps:	3
Percent Similarity:	55.319	Percent Identity:	25.532

```
alignment_block:
US-09-516-493-10 x US-09-031-392-1 ..
```

Align seg 1/1 to: US-09-031-392-1 from: 1 to: 2343

[illegible]

seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-09-299-549-1

```

seq=documentation_block
: Sequence 1, Application US/09299549
: Patent No. 6136547
:
: GENERAL INFORMATION:
:
: APPLICANT: Tartaglia, Louis A.
:
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
: TITLE OF INVENTION: GLUTEX AND USES THEREOF
:
: NUMBER OF SEQUENCES: 10
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
:
: STATE: MA
:
: COUNTRY: USA
: ZIP: 02110-2804
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
:
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows95

```

```

1 SOFTWARE: FASTSEQ for Windows Version 2.0
2
3 CURRENT APPLICATION DATA:
4 APPLICATION NUMBER: US/09/299,549
5 FILING DATE: 26-APR-1999
6
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER: 09/031,392
9 FILING DATE: 26-FEB-1998
10 ATTORNEY/AGENT INFORMATION:
11 NAME: Melkielejohn, Ph.D., Anita L.
12 REGISTRATION NUMBER: 35,283
13 REFERENCE/DOCKET NUMBER: 07334/072002
14
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: 617/542-5070
17 TELEFAX: 617/542-8906
18
19 TELEX: 200154
20
21 INFORMATION FOR SEQ. ID NO.: 1:
22 SEQUENCE CHARACTERISTICS:
23 LENGTH: 2343 base pairs
24 TYPE: nucleic acid
25 STRANDEDNESS: single
26 TOPOLOGY: linear
27
28 MOLECULE TYPE: CDNA
29
30 FEATURE:
31 NAME/KEY: Coding Sequence
32 LOCATION: 73...1761
33
34 US-09-299-549-1

```

```
alignment_scores:
  Quality: 118.50      Length: 141
  Ratio: 1.519        Gaps: 3
Percent Similarity: 55.319  Percent Identity: 25.532
```

```
alignment_block:
US-09-516-493-10 x US-09-299-549-1 .
```

Align seg 1/1 to: US-09-299-549-1 from: 1 to: 2343

```

35  TrrpleuAlaValAlaLysMetCysLeuPheHlaIleAlaCysPheAlaVal  1
   ::::::::::: ::::::::::: :::::::::::
1309  TAcCTAGATGCTGGGCAATTCGTGGCCATATCCCTCTTTCTTCAGTGG  1358
   ::::::::::: ::::::::::: :::::::::::
51  yTrpGlyProIleProCrrPleuLeuMetSerGluIlePheProLeuHisi  68
   ||| |||||:::||||| :::::::::::
1359  GCGAGGCGGATCCCGTTCATCTTGGACTGGAGTCTTCCAGCAATCTC  1408
   ::::::::::: ::::::::::: :::::::::::
68  lElYsglyValAlaIatHnglyAlaCysValLeuHrasnTrpPheMetAla  84
   ::::::::::: ::::::::::: :::::::::::
1409  AGCGGCGCGCTGCCTTCATCTATGGACAGCACCGTCACCTGCCTCCAAC  1458
   ::::::::::: ::::::::::: :::::::::::
85  PheLeuValThirLysGluPheAsnSerIleMetGluIleLeuAtrProTy  101
   ||| ||| ||| ||| ::::::::::: ||| |||
1459  TTTTGCTGTGGGCTCCTCTTCATCCATTCATTCAGAAAGTCTGGACACCTA  1508
   ::::::::::: ::::::::::: :::::::::::
101  rGlyAlaPheTrpLeuHraIleAlaPheCysIleLeuSerValLeuPheTr  118
   | ::::::| ::::| ::::::| ::::::|
1509  C...TGTTTCTGTGCTTCTTCTACAAATTTGTATACAGGTGCTATCTAC  1555
   ::::::::::: ::::::::::: :::::::::::
118  hrLeuThrPheValProGluThrLysGlyAtrThrLeu.....  130
   ::::::::::: ::::::::::: :::::::::::
1556  TGATATTTTGGTGGCTCGTAGACCAAAACAGAACTTATGCAGAAATCAGC  1605
   ::::::::::: ::::::::::: :::::::::::
131  .....GluGlnIleTh  134
   ::::::::::: :::::::::::
1606  CAGGCAATTTCCAAAGAGCAAAAGCATACCCACAGAGAGAAATATCGA  1655
   ::::::::::: :::::::::::
134  rAlaHIsLeuAtrAspGlyAspGlyProLeuSerValThnColysSerProg  151
   ::::| ::::| ::::| ::::|
1656  CTCAGCTGTACATGATGCTCCCTGCTTCTTCTCTTTCACCT...ACTCCGA  1702
   ::::::::::: :::::::::::
151  luleuSerTrpLeuAtrValSer  158
   :::::::::::
1703  ATACAGCTGTGGATTCAGAGCTGCC  1725

```



```
seq_name: /cgn2_6/ptodata/2/ina/5a_COMB.seq:US-08-730-771-1
seq_documentation_block:
: Sequence 1, Application US/08730771
: Patent No. 5780609
: GENERAL INFORMATION:
: APPLICANT: Marshall, Lisa
: APPLICANT: Roshak, Amy
: TITLE OF INVENTION: A NOVEL HUMAN B-CELL SURFACE MOLECULE
: NUMBER OF SEQUENCES: 1
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SmithKline Beecham Corporation
: STREET: 709 Swedeland Road
: CITY: King of Prussia
: STATE: PA
: COUNTRY: USA
: ZIP: 19406-2799
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/730,771
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Han, William T
: REGISTRATION NUMBER: 34,344
: REFERENCE/DOCKET NUMBER: P50380
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-270-5219
: TELEFAX: 610-270-5090
: TELEX:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2775 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE:
: ORIGINAL SOURCE:
: US-08-730-771-1

alignment_scores:
: Quality: 101.00 Length: 176
: Ratio: 1.217 Gaps: 10
: Percent Similarity: 47.159 Percent Identity: 27.273

alignment_block:
US-09-516-493-10 x US-08-730-771-1 ..
Align seg 1/1 to: US-08-730-771-1 from: 1 to: 2775

1 TTPATGProleu**Asn**TTPilleProArGaAlaAlaGlyIlearghi 17
|||||
56 TGGCGCGGCTCTAGACTAGTCGATCCCGCGCTCGCAGCATTCGGCA 105
|||||
17 SGileuValProIleSerAlaGlu...ProAlaAspValHisleuGlyL 33
|||||
106 CGAGGTAACCCACCAAGCAATCTTACCTGTG..... 139
|||||
33 euAlaTrPleuAlaValGlySerMetCysleuPheIleAlaGlyPheAla 49
|||||
140 .....ATGCCGTTTGACCGTCAGCTCTTCTTT..... 166
```

```
50 ValGlyTPGlyProIleProTrPleuLeuMetSerGluIlePheProle 66
|||||
167 .....TGGTGGTGTG..... 178
|||||
66 uHisIleuSGlyValAlaIleThrGlyValCysValIleuThrAsnTrpPheM 83
|||||
179 .....TTTTCGCCGCTGTAAAGTCATCACCCTCTGGGATC 215
|||||
83 etAlaPheleuValThrlySGlyPheAsn..... 92
|||
216 AG...ATGTGCATTCAGAGAACGCCAACAAACATATACTGTGAAT 262
|||||
263 TTAGCTTCAGTGAATCCCTGACACTGTACCAACACAGCAATTTT 312
|||||
106 uThrAlaAlaPheCysIleleu.....Serv 115
|||||
313 GGAATTCAGCTTTAATTTTTCCTACAAATTCACAAATAGAACCTTCAGCA 362
|||||
115 aLeuPheThrIleuThrPheValProGluThrlySGlyArgThr..... 129
|||||
363 GACTCATGATCTTACCTTTTGATTTAAGTTCAGTGCAGATTAAGTGC 412
|||||
130 LeuGluGluIleThrAlaHisleuArGAspGlyAspGlyProleuSerVa 146
|||||
413 ATACATGAAGACACTTTCAAGCCATCATCATTAAGACACTGTGT 462
|||||
146 lThrGlySerProGluLeu.SerTrp 154
|||||
463 AACTGGAATCCCTGATTCATG 488

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-060-208-1
seq_documentation_block:
: Sequence 1, Application US/09060208
: Patent No. 6221629
: GENERAL INFORMATION:
: APPLICANT: Marshall, Lisa
: APPLICANT: Roshak, Amy
: TITLE OF INVENTION: THE DNA SEQUENCE OF HUMAN RP-105
: NUMBER OF SEQUENCES: 1
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Ratner & Prestia
: STREET: P.O. Box 980
: CITY: Valley Forge
: STATE: PA
: COUNTRY: USA
: ZIP: 19482-0980
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/060,208
: FILING DATE: April 15, 1998
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Prestia, Paul F.
: REGISTRATION NUMBER: 23,031
: REFERENCE/DOCKET NUMBER: P50380-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-407-0700
: TELEFAX: 610-407-0701
: TELEX: 846169
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2775 base pairs
```

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-09-060-208-1

alignment_scores:
Quality: 101.00 Length: 176
Ratio: 1.217 Gaps: 10
Percent Similarity: 47.159 Percent Identity: 27.273

alignment_block:
US-09-516-493-10 x US-09-060-208-1 ..

Align seg 1/1 to: US-09-060-208-1 from: 1 to: 2775

```

1 TrratgProleu***Asn***rrlleproargalaiaaglyllearghl 17
|||||
56 TGGCGCGCGCTCTAGAACTAGTGCATCCCGCGCTGCAGGAATTCGGCA 105
17 sgluleuValprolleseeralaglu...prolalspValhlsleuglyL 33
|||||
106 CGAGGGTAACCCACCAAGCAATCTAGCCTGTG..... 139
33 euAlatrleuAlaValglysermetcysleupheIllealaglypheala 49
|||||
140 .....ATGCGTTTgacgtcagctgcttctt..... 166
50 ValglyTrrpglyProilleProtrleuLeumetSerGluIlephrophe 66
|||||
167 .....TGGGTGGTGGCTG..... 178
66 whlslelysglyValAlatrhrGlyValcysValleuthrasntrpphem 83
|||||
179 .....TTTCTGCGCGCTGTAAAGTCAATCCCTCTCTGGGATC 215
83 etAlaphelenuValthrlysgluupheasn..... 92
216 AG..ATGTGCAATGAGAAAGAACCAACAATATTAAGTGTGAAT 262
93 ...SerllemetGluIle.....leuArgProTrrglyAlaphetrrple 106
|||||
263 TTAGGCTTCAGTGAATCCCTGCACACTCTACCAACACAGCAATTTT 312
106 uThrAlaphelhecgyslleleu..... 315
313 GGAATTCAGCTTAAATTTTTCCTCAATTCACATAGACACCTTCGCA 362
115 alLeuphetrrleuthrphelValProgluThrlysglyargThr..... 129
|||||
363 GACCTCATGATCTTACCTTTTGGATTTAAGTAGAGCAATTAAGTGG 412
130 leugluGluIleThrAlahlsleuArgaspglyaspGlyProleuserva 146
|||||
413 ATACATGAGACACCTTTCAAAGCATATCAATTAAGACACACTGTGT 462
146 lThrGlyserProgluleuSerTrr 154
463 AACTGGAATCCCTGATATTCATG 488

```

seq_name: /cgn2_6/ptodata/2/ina/6b_COMB.seq:US-08-818-112-101

seq_documentation_block:
Sequence 101, Application US/08818112
Patent No. 6290969
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.

```

APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Vedicik, Thomas S.
APPLICANT: Twardzik, Daniel R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESS: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,112
FILING DATE: 13-MAR-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 500 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-818-112-101

```

alignment_scores:
Quality: 97.00 Length: 22
Ratio: 5.105 Gaps: 0
Percent Similarity: 86.364 Percent Identity: 77.273

alignment_block:
US-09-516-493-10 x US-08-818-112-101/rev ..

Align seg 1/1 to reverse of: US-08-818-112-101 from: 1 to: 500

```

1 TrratgProleu***Asn***rrlleproargalaiaaglyllearghl 17
|||||
487 TGGCGCGCGCTCTAGAACTAGTGCATCCCGCGCTGCAGGAATTCGGCA 438
17 sgluleuValprolle 22
|||||
437 CGAGATGCGCGCGCTT 422

```

seq_name: /cgn2_6/ptodata/2/ina/5a_COMB.seq:US-07-968-971A-10

seq_documentation_block:
Sequence 10, Application US/07968971A
Patent No. 5455167
GENERAL INFORMATION:
APPLICANT: Toni A. Voelker
APPLICANT: Huw Maelor Davies
TITLE OF INVENTION: Medium-Chain Thioesterases
TITLE OF INVENTION: In Plants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis

STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB
OPERATING SYSTEM: Macintosh 6.0.7
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/968,971A
FILING DATE: 19921030
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/824,247
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/04332
FILING DATE: 21-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
REFERENCE/DOCKET NUMBER: CGNE 88
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-07-968-971A-10

alignment_scores:
Quality: 97.00 Length: 21
Ratio: 5.389 Gaps: 0
Percent Similarity: 85.714 Percent Identity: 85.714

alignment_block:

US-09-516-493-10 x US-07-968-971A-10 ..

Align seg 1/1 to: US-07-968-971A-10 from: 1 to: 1601

1 TTPATGProLeu**Asn**TriPleProAGAlaIaGlyIleArgH 17
||||| 111 ||| ||||||||||||||||||||
7 TGGCGCGCGCTTAGACTAGTGCATCCCGCGGCTGCAGCAATTGCGCA 56
17 sGluLeuValPro 21
||||| 111
57 CGAGCTTCTCTCCC 69

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-424-406-6

seq_documentation_block:

Sequence 6, Application US/08424406
Patent No. 567997
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
APPLICANT: Davies, Huw Maelor
APPLICANT: Knutzen, Deborah S.

TITLE OF INVENTION: Medium Chain Thioesterases in Plants
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1(a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,406
FILING DATE: 26 April, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10814
FILING DATE: 29 October, 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US93 07/968, 971
FILING DATE: 30 October, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 88-105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-424-406-6

alignment_scores:
Quality: 97.00 Length: 21
Ratio: 5.389 Gaps: 0
Percent Similarity: 85.714 Percent Identity: 85.714

alignment_block:

US-09-516-493-10 x US-08-424-406-6 ..

Align seg 1/1 to: US-08-424-406-6 from: 1 to: 1601

1 TTPATGProLeu**Asn**TriPleProAGAlaIaGlyIleArgH 17
||||| 111 ||| ||||||||||||||||||||
7 TGGCGCGCGCTTAGACTAGTGCATCCCGCGGCTGCAGCAATTGCGCA 56
17 sGluLeuValPro 21
||||| 111
57 CGAGCTTCTCTCCC 69

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-472-934-5

seq_documentation_block:

Sequence 5, Application US/08472934
Patent No. 575346
GENERAL INFORMATION:
APPLICANT: JOHNSON, GARY L.
TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
RESPONSIVENESS TO EXTERNAL SIGNALS

130 CGGCCGGTCGAGTGGCCCGCGAGCGGGTTGGCCGAGCTTGG 176
seq_name: /cgn2_6/plodata/2/ina/5B_COMB.seq:US-08-461-146C-5
seq_documentation_block:
: Sequence 5, Application US/08461146C
: Patent No. 5981265
: GENERAL INFORMATION:
: APPLICANT: JOHNSON, GARY L.
: TITLE OF INVENTION: METHODS FOR REGULATING MEKK PROTEIN ACTIVITY
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lahive and Cockfield
: STREET: 60 State Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/461,146C
: FILING DATE: 05-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/354,516
: FILING DATE: 21-FEB-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/049,254
: FILING DATE: 15-APR-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/323,460
: FILING DATE: 14-OCT-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US94/11690
: FILING DATE: 14-OCT-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US94/04178
: FILING DATE: 15-APR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: KARA, Catherine J.
: REGISTRATION NUMBER: P41,106
: REFERENCE/DOCKET NUMBER: CPI-004CNI3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 227-5941
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3089 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 400..2280
: US-08-461-146C-5

alignment_scores:
: Quality: 96.00 Length: 49
: Ratio: 3.429 Gaps: 2
Percent Similarity: 57.143 Percent Identity: 46.939

alignment_block:
US-09-516-493-10 x US-08-461-146C-5 ..
Align seg 1/1 to: US-08-461-146C-5 from: 1 to: 3089
1 TTPATgProLeu***Asn***TtpIleProARgAlaIaGlyIleArgHi 17

||||||||| ||| |||||||||||||||||||
30 TGGCGGCGCTCTAGAACTAGTGATCCCGCGGCTGCGAGAAATTCGGCA 79
17 s...GlueValProIleSerAlaIu..... 25
| ||| |||||||:
80 CGAGAACAGATGGCGGTCGAGAGCTTTCTGTGACTTCAGACTCGCAG 129
26ProIaAspValHisLeuGlyLeuAlaTrp 35
|||||: |||||
130 CGGCCGGTCGAGTGGCCCGCGAGCGGGTTGGCCGAGCTTGG 176
seq_name: /cgn2_6/plodata/2/ina/6A_COMB.seq:US-08-461-145C-5
seq_documentation_block:
: Sequence 5, Application US/08461145C
: Patent No. 6074861
: GENERAL INFORMATION:
: APPLICANT: JOHNSON, GARY L.
: TITLE OF INVENTION: NOVEL MEKK PROTEINS
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lahive and Cockfield, LLP
: STREET: 60 State Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/461,145C
: FILING DATE: 5-JUNE-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/049,254
: FILING DATE: 11-APR-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/323,460
: FILING DATE: 14-OCT-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US94/11690
: FILING DATE: 14-OCT-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US94/04178
: FILING DATE: 15-APR-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/354,516
: FILING DATE: 21-FEB-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Kara, Catherine J.
: REGISTRATION NUMBER: P-41,106
: REFERENCE/DOCKET NUMBER: CPI-004CNI
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 227-5941
: INFORMATION FOR SRO ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3089 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 400..2280
: US-08-461-145C-5

alignment_scores:
: Quality: 96.00 Length: 49


```

286 GGGAGACCATCCCTGGCTCCTCATGTCAGAGATCTTCCCTGCAATGTC 335
69  LysGlyValAlaThrGlyValCysValLeuThrAsnTrpPheMetAlaPh 85
|||||
336 AAGGCTGGCTACCGGATCTGTCTCTCACCAACGTGCTTACGCGCTT 385
85  eLeuValThrLysGluPheAsnSerIleMetGluIleLeuArgProTyrG 102
|||||
386 TCTAGTACCAAGAGATTACACGACGTCATGCAATGCTCAGACCTACG 435
102  LysAlaPheTrpLeuThrAlaAlaPheCysIleLeuSerValLeuPheThr 118
|||||
436 GTGGCTTGTGGCTACCGCTGTCTTCTGG...CTTCAAGTCTCCTATTGC 482
119  LeuThr...PheValProGluThrLysGlyArgThrLeuGluGluIleThr 134
|||||
483 CACACTGACCGCTGCCCTGAGACTAAAGCAGGACTCTGGACACAGTCGA 532
134  rAlaHisLeuArgAspGlyAspGlyProLeuSerValThrGlySerProG 151
|||||
533 CAGGCCATTTCTGAGGAGGAGATGACAGACCTTCTGTGTGGAGGCCCG 582
151  IuLeuSerTrpLeuArgValSerLysGlyValGluTrp 163
|||||
583 AGCTGAGCTGGCTTTGGCTTTCAGAGAGGTGAGTGG 620

seq_name: gb_est2:H34451
seq_documentation_block:
LOCUS      H34451          319 bp      mRNA          EST          13-MAR-1998
DEFINITION ESF111406 Rat PC-12 cells, NGF-treated (9 days) Rattus sp. cDNA 5'
end similar to Glucose transporter, mRNA sequence..
ACCESSION  H34451
VERSION    H34451.1  GI:979868
KEYWORDS   EST.
SOURCE     Rattus sp.
ORGANISM   Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 319)
AUTHORS   Lee,N.H., Weinstein,K.G., Kirkness,E.F., Earle-Hughes,J.A., Fuldner
            ,R.A., Marmaras,S., Glodde,A., Gocayne,J.D., Adams,M.D., Kerlavage
            ,A.R., Fraser,C.M. and Venter,J.C.
            Comparative expressed-sequence-tag analysis of differential gene
            expression profiles in PC-12 cells before and after nerve growth
            factor treatment
            Proc. Natl. Acad. Sci. U.S.A. 92, 8303-8307 (1995)
JOURNAL   95396786
COMMENT    Contact: Lee, NH
            The Institute for Genomic Research
            9712, Medical Center Drive, Rockville, MD 20850, USA
            Tel: (301)-838-5529
            Fax: (301)-838-0208
            Email: nhlee@iglr.org
            For clone availability please contact the TIGR Database
            (cdinfo@tdb.iglr.org)
            Seq primer: M13 Reverse.

FEATURES
            source
                Location/Qualifiers
                    1..319
                        /organism="Rattus sp."
                        /db_xref="ATCC (inhost):2005510"
                        /db_xref="taxon:10118"
                        /clone_lib="Rat PC-12 cells, NGF-treated (9 days)"
                        /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                        XhoI; poly(A)+ RNA was purified from 9-day NGF treated
                        PC12 cells. cDNA was constructed using an oligo-dn primer
                        and directionally cloned using the lambda ZAP II Vector
                        Kit by Stratagene"
BASE COUNT  48 a 102 c 80 g 88 t 1 others
ORIGIN

```

```

alignment_scores:
    quality: 561.00      Length: 106
    Ratio: 5.343
    Percent Similarity: 99.057      Percent Identity: 99.057

alignment_block:
US-09-516-493-10 x H34451
Align seq 1/1 to: H34451 from: 1 to: 319

19  LeuValProIleSerAlaGluProAlaAspValHisLeuGlyLeuAlaTr 35
|||||
1  CTGGTGGCCATCTCCGACAGGCTGCTGATGTTCACCTGGGCTGGCTG 50
35  PleuAlaValAlGlySerMetCysLeuPheIleAlaGlyPheAlaValAlG 52
|||||
51  GCTGGCTGTAGGACGATGTCCTCTTCAATCCGTGGTTTGGACGTAGGCT 100
52  rPGLYProIleProTrpPLeuLeuMetSerGluIlePheProLeuHisIle 68
|||||
101  GGGAGCCATCCCTGGCTCCTCATGTCAGAGATCTTCCCTGCACATC 150
69  LysGlyValAlaThrGlyValCysValLeuThrAsnTrpPheMetAlaPh 85
|||||
151  AAGGCTGTGGCTACCGGCTGTCTGTCTCACCACACTGTTATGAGCTT 200
85  eLeuValThrLysGluPheAsnSerIleMetGluIleLeuArgProTyrG 102
|||||
201  TTTGGTGACCAAGAGTTTACAGCATCATGAGATTCACAGACCTACG 250
102  lYAlaPheTrpLeuThrAlaAlaPheCysIleLeuSerValLeuPheThr 118
|||||
251  GCGCTTTCGGCTCACCGCTGCTCTGTATCTCAGCGTCTTTTCACG 300
119  LeuThrPheValProGlu 124
|||||
301  CTCACCTTTGTCCCTGAG 318

seq_name: gb_est2:BG749509
seq_documentation_block:
LOCUS      BG749509       949 bp      mRNA          EST          15-MAY-2001
DEFINITION 602707648F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:484411 5',
mRNA sequence.
ACCESSION  BG749509
VERSION    BG749509.1  GI:14060162
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 949)
AUTHORS   NIH-MGC http://mgc.ncl.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgabbs@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L16M1681 row: p column: 12
            High quality sequence stop: 745.

FEATURES
            source
                Location/Qualifiers
                    1..949
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:484411"
                        /clone_lib="NIH_MGC_43"
                        /tissue_type="normal pigmented retinal epithelium"

```



```

/lab_host="DH10B (phage-resistant)"
/Note="Organ: eye; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library. |"
BASE COUNT      160 a      302 c      264 g      223 t
ORIGIN
alignment_scores:
  Quality: 552.00      Length: 125
  Ratio: 4.678      Gaps: 0
Percent Similarity: 94.400      Percent Identity: 80.000
alignment_block:
US-09-516-493-10 x BG749509 ..
Align seg 1/1 to: BG749509 from: 1 to: 949
21 ProIleSerAlaGluProAlaAspValHisLeuGlyLeuAlaTrpLeuAl 37
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
68 CCTGCTCTCCACAGCGCTTGATGCCAGCGCGCGCTGGCTGGCTGCG 117
37 aValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyP 54
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
118 CGTGCGCAGCATGTGCTCTTCATGCGCGGCTTGGGGTGGCTGGGCG 167
54 rolIeProTrpLeuLeuMetSerGluIlePheProLeuHisIleLysGly 70
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
168 CCATGCCCTGGCTCTCATGTGACAGATCTGCCCTGCATGTCACAGGCG 217
71 ValAlaThrGlyValCysValLeuThrAsnTrpPheMetAlaPheLeuVa 87
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
218 GTGGCGACGAGCATCTGCTGCTCACCAACTGGCGTCATGGCGCTTCTCG 267
87 lThrLysGluPheAsnSerIleMetGluIleLeuAlaTrpGlyAlaP 104
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
268 GACCACAGAGTTCACAGCGCTCATGAGAGCTCTCAGCGCCATGAGAGCT 317
104 heTrpLeuThrAlaAlaPheCysIleLeuSerValLeuPheThrLeuThr 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
318 TCTGGCTTGCCCTCGCTTTCGATCTTCAGTCTTTCACATTTGCTC 367
121 PheValProGluThrLysGlyArgThrLeuGluGlnIleThrAlaHisLe 137
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
368 TGTGTCCCTGAACCTAAAGGAAGACTCTGGAACAATACAGACCCATT 417
137 uArgAspLysAspGlyProLeuSer 145
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
418 TGAAGGGGAGATGACAGCCACTCACT 442
seq_name: gb_est1:BE231636
seq_documentation_block:
LOCUS      BE231636      546 bp      mRNA      EST      10-JUL-2000
DEFINITION 136437 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION  BE231636
VERSION    BE231636.1 GI:9016354
KEYWORDS   EST.
SOURCE     pig.
ORGANISM  Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
REFERENCE  1 (bases 1 to 546)
AUTHORS   Fahrkenrugg,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grose,W.M., Bennett,G.A., Laegreid,W.M.
and Keele,J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine

```

```

JOURNAL      Unpublished (2000)
COMMENT      Contact: Smith TPL
              USDA, ARS, US Meat Animal Research Center
              PO Box 166, Clay Center, NE 68933-0166, USA
              Tel: 402 762 4366
              Fax: 402 762 4390
              Email: smith@email.marc.usda.gov
              Single pass sequencing. Bases called and alt-trimmed with phred
              v0.980904.e. Vector identified by cross_match with the -mnscore 18
              and -mismatch 12 options.
              PCR primers:
              FORWARD: AGCAACACGCTATGACCAT
              BACKWARD: GTTTCCAGTCACGACG
              Plate: 70 row: A column: 21
              Seq primer: ATTAGTCACACTATAG.
FEATURES
  source
    1..546
    /organism="Sus scrofa"
    /db_xref="taxon:9823"
    /clone_lib="MARC 1P1G"
    /tissue_type="pooled"
    /lab_host="DH10B"
    /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
    Library made from pooled tissue from day 11, 13, 15, 20,
    and 30 embryos."
BASE COUNT      76 a      182 c      162 g      124 t      2 others
ORIGIN
alignment_scores:
  Quality: 548.00      Length: 140
  Ratio: 4.349      Gaps: 3
Percent Similarity: 90.000      Percent Identity: 75.714
alignment_block:
US-09-516-493-10 x BE231636 ..
Align seg 1/1 to: BE231636 from: 1 to: 546
25 GluProAlaAspValHisLeuGlyLeuAlaTrpLeuAlaValGlySerMe 41
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2  GAGCGCGCGGATGCCAGCGGTGGGCTGCTGCTGGCGGTGGCAGCGCT 51
41 lCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProLeuTrpL 58
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
52 GTGCCCTTTCATTTGCTGGCTTGCGCTGGCGTGGGGCCCATCCCTGGC 101
58 euLeuMetSerGluIlePheProLeuHisIleLysGlyValAlaThrGly 74
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
102 TCCATGATGTGAGATCTTCCCTGACGTCACAGGCGGTGGCCACCGGC 151
75 ValCysValLeuThrAsnTrpPheMetAlaPheLeuValThrLysGluPh 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
152 GTCTGTGCTCTACCAACTGGCTCATGGCGCTTCTGTGTGACAAAGAGTT 201
91 eaSnSerIleMetGluIleLeuArgProTrpGlyAlaPheTrpLeuThr 108
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
202 CAGCAGCGCTCATGAGAGTGTCTCAGGCCCTTACGGTCTTGTGGCTGCT 251
108 lAlaPheCysIleLeuSerValLeuPheThrLeuThrPheValProGlu 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
252 CCGCTTTCGATCTTCAGTGTCTTTCACATCTGAGCGCTGTCCCTGAA 301
125 ThrLysGlyArgThrLeuGluGlnIleThrAlaHisLeuArgAspGlyA 141
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
302 ACCAAGGAGAGACTCTGAGCAGATCACAGCCCATTTTGAAG...GGCG 348
141 spGlyProLeuSerValThrGlySerProGluLeuSerTrpLeuArgVal 157
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
349 ATGACAGCGCTTCCCGGAGAGTGGCCACCCCTTANCTGACACTGGCTT 398
158 ...SerLysGlyValGlu 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```



```

47 yPheAlaValGlyTrpGlyProIleProTrpLeuLeuSerGlyIleP 64
   |||
56 CTT.GCGGTGGCTGGGGGCCATCCCTGGCTCCTCATGTACAGACT 104
64 hePheLeuHisIleLysGlyAlaIleThrGlyValCysValLeuThrAsn 80
   |||
105 TCCCTGTGCAATGCAGGCGCTGGCGACAGGCATCTGGCTCCACCAAC 154
81 TrpPheMetAlaPheLeuValThrLysGlyIlePheAsnSerIle.MetGlyI 97
   |||
155 TGGCTCATGGCTTTCTCTGTACCAAGAGATTCCAGCAGCTCTATGAGAG 204
97 IeLeuArgProTyrGlyAlaPheTrpLeuThrAlaIlePheCysIleLeu 113
   |||
205 TCTCAGAGCCCTATGAGAGCTTCTGGCTGGCTGCGCTTCTGCATCTTC 254
114 SerValLeuPheThrLeuThrPheValProGlyThrLysGlyArgThrLe 130
   |||
255 AGTGTCTTTTCACTTGTCTGT.GTCCCTGAACCTAAAGAAAGACATCT 303
130 uGluGlnIleThrAlaHisLeuArgAspGlyAsp 141
   |||
304 GGAACAATTCACAGCCCATTTGAGGGCGGATGAC 337
seq_name: gb_est2:BE910478

```

```

seq_documentation_block: 934 bp mRNA EST 20-OCT-2000
LOCUS BE910478 mRNA
DEFINITION 601501005F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3903048 5',
            mRNA sequence.
ACCESSION BE910478
VERSION BE910478.1 GI:10407108
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 934)
            NIH-MGC http://mgs.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            JOURNAL Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LAM9706 row: 1 column: 01
            High quality sequence stop: 710.
FEATURES
    source
        1..934
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:3903048"
            /clone_lib="NIH_MGC_70"
            /tissue_type="epithelioid carcinoma"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: pancreas; Vector: pCMV-Sport6; Site_1: NotI;
            Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
            Average insert size 1.1 kb. Library constructed by Life
            Technologies."

```

```

BASE COUNT 143 a 291 c 287 g 213 t
ORIGIN
alignment_scores:
    Quality: 464.00 Length: 200
    Ratio: 3.742 Caps: 4
Percent Similarity: 62.000 Percent Identity: 50.000

```

```

alignment_block:
us-09-516-493-10 x BE910478
Align seg 1/1 to: BE910478 from: 1 to: 934

```

```

14 GlyIleArgHisGlyLeuVal.....ProIleSerAlaGlyIleP 20
   |||
279 GGTGTTCACAGCAGACTGCTTGGCGGCGCTACTTCAAGCTGACCCAGTGG 328
21 .....
329 CCCTGGCAACTCTCTGCACAGTGGCCCATCTGGCGCTTCTGTGCACAC 378
26 roAlAspValHisLeuGlyLeuAlaIleThrPheValAlaGlySerMetCys 42
   |||
379 CTGTTGATGCCAGCGTGGGGCTGGCTGGCTGGCGGCGGACATGTGC 428
43 LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeu 59
   |||
429 CTCTTCATGCGCGGCTTGGCGGTGGCTGGCGGCCCATCCCTGGCTCCT 478
59 uMetSerGlyIlePheProLeuHisIleLysGlyValAlaThrGlyValC 76
   |||
479 CATGTTCAGAGATCTTCCCTCTGCATGTCAAGGCGCTGGCGAAGGCATCT 528
76 ySValLeuThrAsnTrpPheMetAlaPheLeuValThrLysGlyIlePheAsn 92
   |||
529 GCGTCTTCACCACTAGCGCTCATAGGCTTCTGTGACCAAGAGATTTCAC 578
93 SerIleMetGlyIleLeuArgProTyrGlyAlaIleThrPheThrAlaAl 109
   |||
579 AGCTTCATGAGAGCTCTCAGGCGCTATGAGGCTTCTGGCTTGCCTCCT 628
109 aPheCysIleLeuSerValLeuPheThrLeuThrPheValProGlyThrL 126
   |||
629 TTCTGCATCTTCAGTGGTCTTTTACCTTGTCTGTGCTGTAAGAACTA 678
126 ySGlyArgThrLeuGluGlnIleThrAlaHisLeuArgAsp..... 139
   |||
679 AGCAAGAGA...CTCTGGAACATTCACAGCATTTTGAGGGCGGATGACC 725
139 .....
726 GCCATCATAGCGGTTGGCGACAGCTGTGATCCAGTGGGCGCAGCCACA 775
140 .....GlyAspGlyProLeuSer 145
776 GCCCTGCTGCCAGAGACAGATCAGCTGGCTTGTGGGCGACCCCTCGGC 825
146 ValThrGlySerProGlyLeuSerTrpLeuArgValSerLysGlyVal 161
   |||
826 GTCTCCCTAACCGGCTGAGAGCGCTGTGCTGTGCTGGCGGACTC 873

```

```

seq_name: gb_est2:BG672321
seq_documentation_block: 538 bp mRNA EST 30-APR-2001
LOCUS BG672321 mRNA
DEFINITION DRNCPD10 Rat DRG library Rattus norvegicus cDNA clone DRNCPD10 5',
            mRNA sequence.
ACCESSION BG672321
VERSION BG672321.1 GI:13894420
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE 1 (bases 1 to 538)
AUTHORS Xiao,H.S., Han,Z.G., Zhang,F.X., Huang,Q.H., Lu,Y.J., Bao,L., Fu,G.,
            Guo,C., Yan,Q., Jin,S.X., Zhu,Z.D., Xu,X.R., Li,N.G., Chen,Z. and
            Zhang,X.
TITLE Distinct gene expression profiles of rat dorsal root ganglion

```

JOURNAL
COMMENT

Induced by peripheral nerve axotomy
Unpublished (2001)
Contact: Zhang Xu
Laboratory of Sensory System
Institute of Neuroscience
320 Yue Yang Road, Shanghai 200031, P. R. China
Tel: 86-21-64748700-121
Fax: 86-21-64713446
Email: xu.zhang@ion.ac.cn
This clone is also available at Chinese National Human Genome
Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park,
Pudong New Area, P. R. China. Please contact with Zhang Xu
(xu.zhang@ion.ac.cn) or Han Zeyang (hanzy@chgc.sh.cn)
PCR Primers
FORWARD: T3
BACKWARD: T7
Seq primer: T3
POLY-A=No.

FEATURES
source

1. 538
Location/Qualifiers
/organism="Rattus norvegicus"
/strain="Sprague Dawley"
/db_xref="taxon:10116"
/clone="DRNCPD10"
/clone_lib="Rat DRG Library"
/sex="male"
/tissue_type="dorsal root ganglion"
/dev_stage="adult"
BASE COUNT 118 a 146 c 138 g 136 t
ORIGIN

alignment_scores:
Quality: 459.00 Length: 113
Ratio: 4.413 Gaps: 1
Percent Similarity: 92.035 Percent Identity: 82.301

alignment_block:
US-09-516-493-10 x BG672321 ..

Align seg 1/1 to: BG672321 from: 1 to: 538

```

53 GlyProIleProTrrPleuLeuMetSerGluIlePheProLeuHisIleLeu 69
|||||
5  GGACCCATTCCTCGCTCTTATTCAAGAAGCTTCCTTCACAAATGA 54
|||||
69  sGlyValAlaThrGlyValCysValLeuThrAsnTrpPheMetAlaPheLeu 86
|||||
55  GGGGGGGGTACCGGGGCTCGGCTCCACCAACTGGTTTCATGGCGCTTTC 104
|||||
86  euValThrLysGluPheAsnSerIleMetGluIleLeuArgProTyrGly 102
|||||
105  TGGGACCAAGAGTTTAACAGCATCATGAGAACTTTAAACCTACAGGCG 154
|||||
103  AlaPheTrpLeuThrAlaAlaPheCysIleLeuSerValLeuPheThrLe 119
|||||
155  GCCTTTGGCTCAACGCTGCTTTGTATTCTTACGCGCTTTCACGCT 204
|||||
119  uThrPheValProGluThrLysGlyArgThrLeuGluGlnIleThrAlaH 136
|||||
205  CACCTTTCCTTCCTTGACTAAAGGACAGCTTTGGACAAATACACAGGCG 254
|||||
136  Is.LeuArgAspGlyAspGlyProLeuSerValThrGlySerProGluLe 152
|||||
255  ATTGTGAGGAGGAGTGAGCGACCCCTT.TGTGTGACGGGCAACCCCTAGCT 303
|||||
152  uSerTrpLeuArgValSerLysGlyValGluTrpPro 164
|||||
304  GAGCTGGCTTCGGTTTCAAAAGAGAGTGAGTGGCGCT 340
|||||

```

seq_name: gb_est1:AW248655
seq_documentation_block:

LOCUS AW248655 496 bp mRNA EST 07-JAN-2000
DEFINITION 2820759.3prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820759 3',
mRNA sequence.
ACCESSION AM248655
VERSION AM248655.1 GI:6591648
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS NIH-MGC
TITLE NIH-MGC
JOURNAL http://mgc.nci.nih.gov/.
COMMENT Unpublished (1999)
Other ESTs: 2820759.5prime
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov

Tissue Procurement: DCTD/DTF cDNA Library Preparation: Ling
Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
project clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bdr/image/image.html Base Calling / Quality
Scores: PHRED from University of Washington Genome Center. Vector
Trimming: cross_match from University of Washington Genome Center
PHRAP suite. Poly-T identification: patmatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:
http://www.genome.washington.edu polyadenylation: Based upon the
presence of a XhoI site followed by a run of 14 or more T residues
at the beginning of the sequence. This cDNA insert was
polyadenylated.
Plate: LLCMS row: A column: 16
High quality sequence stop: 415.

FEATURES
source

1. 496
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2820759"
/clone_lib="NIH_MGC-7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site.1: XhoI; Site.2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GCCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 116 a 145 c 142 g 92 t 1 others
ORIGIN

alignment_scores:
Quality: 365.00 Length: 78
Ratio: 4.803 Gaps: 0
Percent Similarity: 97.436 Percent Identity: 82.051

alignment_block:
US-09-516-493-10 x AW248655/rev ..

Align seg 1/1 to reverse of: AW248655 from: 1 to: 496

```

21  ProIleSerAlaGluProAlaSpValHisLeuGlyLeuAlaTrpLeuAl 37
|||||
467  CCTGTCTNTGACAGCGCTTTGATGCGACGCTGGCGCTGGCGTGGC 418
|||||
37  aValGlySerMetCysLeuPheIleAlaGlyPheAlaValAlaGlyTrpGlyP 54
|||||
417  CCGGGGACAGCATGTGCTTCATCGCCGCTTGGCGGTGGCGGCGC 368
|||||
54  rolleProTrrPleuLeuMetSerGluIlePheProLeuHisIleLysGly 70
|||||

```



```

|||||
52 CGCTGCGCTCCACCACTGGCTTCATGGCTTCAGTACCAAGAGT 101
91 heanserilemetgluileuargprotyrallyalaphetrplethr 107
102 TCAGCAGCGTCATGAGAGCCCTCGGCCCTACGGCGGCTTCGGCTGCC 151
108 AlaAlaPheCysIleLeuSerValIleuPheThrIleuValProgl 124
152 TCTGCGCTTCGATCTCGCGCTTCCTTTCATCTTCTGCGCTCGA 201
124 uthrlysglyargthrleugluinlethrila 135
202 ACCAAGGGGAGAGACTCTGGAACAATCAGAGCC 235

seq_name: gb_est2:BF742266

seq_documentation_block:
LOCUS BF742266 327 bp mRNA EST 10-JAN-2001
DEFINITION RCI-BTN0409-021000-012-e12 BTN0409 Homo sapiens cDNA, mRNA
sequence.
ACCESSION BF742266
VERSION BF742266.1 GI:12069046
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 327)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.,
Nagai, M. A., de Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F.,
Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H.,
Brunstein, J., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare,
M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and
Simpson, A. J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RCI&t2=RCI-BTN0409
-021000-012-e12&t3=2000-10-02&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
Location/Qualifiers
1..327
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BTN0409"
/dev_stage="Adult"
/note="Organ: breast; Vector: puc18; Site_1: Sma1; Site_2:
Sma1; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 82 a 91 c 108 g 46 t
ORIGIN

```

alignment_scores: Quality: 348.50 Length: 84

Ratio: 4.526 Gaps: 3
Percent Similarity: 91.667 Percent Identity: 78.571
alignment_block:
US-09-516-493-10 x BF742266/rev ..
Align seq 1/1 to reverse of: BF742266 from: 1 to: 327

```

21 ProIleSerAlaGluProAlaAspValHisLeu.....GlyLeuAlaTr 35
|||||
297 CGTGTCTCTCAGACGCGCTGTGATGCCAGCGTGGGAGCATGGCGGTGATG 248
35 pleuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGly 52
247 G...CCAGTGGCGAGCATGTGGCTTCATCGCGCGGCTTCGGTGGGCT 201
52 rpglyProIleProIleProIleuMetSerGluIlePheProLeuHisIle 68
200 GGGGGCCCATCCCTGGCTCTCATGTGAGAGATCTTCCTCATGTC 151
69 LysGlyValAlaThrGlyValCysValIleuThrAsnTrpPheMetAlaPh 85
150 AAGGCGTGGCGACAGGATCTGGCTCTCACCAACTGCGTCATGGCTT 101
85 eleuValThr.LysGluPheAsnSerIleMetGluIleLeuArgProTyr 101
100 TCTGCTGACACAGAGAGTTCAGAGCGCTCATGGAGTCTTCAGGCCCTAT 51

seq_name: gb_est2:BF140667

seq_documentation_block:
LOCUS BF140667 917 bp mRNA EST 24-OCT-2000
DEFINITION 601786917F1 NCL_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4014605 5',
mRNA sequence.
ACCESSION BF140667
VERSION BF140667.1 GI:10979707
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 917)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@rs-remail.nih.gov
tissue procurement: Gilbert Smith, Ph.D.
cDNA library preparation: Life Technologies, Inc.
cDNA library arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM9259 row: a column: 06
High quality sequence stop: 605.
Location/Qualifiers
1..917
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_image="IMAGE:4014605"
/clone_lib="NCL_CGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; transgenic model MNV-1, expression driven by
MNV-LTR enhancer; cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT 163 a 263 c 277 g 214 t
ORIGIN

```

```
alignment_scores:
  Quality: 323.00      Length: 77
  Ratio: 4.307        Gaps: 1
  Percent Similarity: 97.403      Percent Identity: 83.117

alignment_block:
  US-09-516-493-10 x BF140667 ..

Align seg 1/1 to: BF140667 from: 1 to: 917

19 LeuValProIleSerAlaGluProAlaAspValHisLeuGlyLeuAlaTr 35
|||||.....|.....|.....|.....|.....|.....|.....|
412 CTGGGCCCATCGCGCGGACCTGTGATGTCCAAAGTGGGAGCTGGCTG 461
35 PleuAlaValGlySerMetCysLeuPhe11AlaGlyPheAlaValGlyT 52
|||||.....|.....|.....|.....|.....|.....|.....|
462 GCTGGCTTAGGCACATGCTGCTTCATTCGTGGCTTTGGCGGGCT 511
52 rpgIyProIleProTrpleuLeuMetSerGluIlePheProLeuHisIle 68
|||||.....|.....|.....|.....|.....|.....|.....|
512 GGGGACCATCTCCCTGGCT. CTCATGTGAGAGATCTCCCTCGATGTC 560
69 LysGlyValAlaThrGlyValCysValLeuThrAsnTrpPheMetAlaPh 85
|||||.....|.....|.....|.....|.....|.....|.....|
561 AAGGTGTGGCTACCGGCGCTGTCTCTCCACCACTGTCATGTGCTT 610
85 eLeuValThr.....LysGluPheAsnSer 93
|||||.....|.....|.....|.....|.....|.....|.....|
611 TTCTCTTAAGTTGACCAAAAGAGTTCAAGCA 641

seq_name: gb_est2:BG303512

seq_documentation_block:
LOCUS BG303512 295 bp mRNA EST 23-FEB-2001
DEFINITION F156g08.x1 Sugano Kawakami zebrafish DNA Danio rerio cDNA clone
3817934 3' similar to TR:Q9W610 Q9W610 CG8234 PROTEIN.; mRNA
sequence.
ACCESSION BG303512
VERSION BG303512.1 GI:13101039
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Rasbora; Danio.
1 (bases 1 to 295)
REFERENCE
  Sugano,S., Kawakami,K., Johnson,S., Li,F., Marra,M., Eddy,S.,
  Hillier,L., Clifton,S., Allen,M., Gibbons,M., Jost,S., Kucaba,T.,
  Martin,J., Pape,D., Steptoe,M., Underwood,K., Treisling,B., Riltter
  ,E., Bowers,Y., Wylie,T., Waterston,R. and Wilson,R.
  Washu zebrafish EST Project 1999
  Unpublished (1999)
  Contact: S.L. Johnson
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@wustl.edu
  Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
  Sequencing by: Washington University Genome Sequencing Center Clone
  distribution information can be found through the I.M.A.G.E.
  Consortium/LLNL, send email to: info@image.llnl.gov
  Trace considered overall poor quality
  Seq primer: T7 from Gibco
  High quality sequence stop: 1.
  Location/Qualifiers
    1..295
    /organism="Danio rerio"
    /strain="AB"
    /db_xref="taxon:7955"
    /clone="3817934"
    /clone_lib="Sugano Kawakami zebrafish DNA"
```

```
/sex="mixed (one male and one female, including
unfertilized eggs)"
/dev_stage="adult"
/lab_host="DH10B (phage resistant)"
/note="Vector: pME18S-FL3; Site_1: DraIII (CACGCTGTG);
Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed
with an oligo(dT) primer [ATCGGCCCTTTTCTTTTCTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[GTGTGGCCCTACTGG], digested and cloned into distinct DraIII
sites of the pME18S-FL3 vector (5' site CACGTGTG, 3' site
CACCATGTG). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed by Dr. Sumio Sugano
(University of Tokyo Institute of Medical Science) and
kindly donated by Dr. Koichi Kawakami. Custom primers for
sequencing: 5' end primer CTGTGCTCTAAAGCTGGC and 3' end
primer CGACCTGACGCTGACGACA."

BASE COUNT 87 a 82 c 72 g 54 t
ORIGIN

alignment_scores:
  Quality: 318.00      Length: 96
  Ratio: 4.025        Gaps: 0
  Percent Similarity: 82.292      Percent Identity: 57.292

alignment_block:
  US-09-516-493-10 x BG303512/rev ..

Align seg 1/1 to reverse of: BG303512 from: 1 to: 295

33 LeuAlaTrpLeuAlaValGlySerMetCysLeuPhe11AlaGlyPheAl 49
|||||.....|.....|.....|.....|.....|.....|.....|
289 CTGGCCTGGTTGCTCTGTGATGACATGGGCTTTTCATTCGACGATTTGC 240
49 aValAlaTrpGlyProIleProTrpleuLeuMetSerGluIlePheProL 66
|||||.....|.....|.....|.....|.....|.....|.....|
239 TATGTGTGGGCTCTACTGCTGCTGTGATGTACAGAGATCTGCCCA 190
66 eHisIleLeuGlyValAlaThrGlyValCysValLeuThrAsnTrpPhe 82
|||||.....|.....|.....|.....|.....|.....|.....|
189 CGGGGTGAGGGGATTAGGACATGCTGTGTGCTTACTTACTTACGTGACC 140
83 MetAlaPheLeuValTrpLysGluPheAsnSerIleMetGluIleLeuAr 99
|||||.....|.....|.....|.....|.....|.....|.....|
139 TGTGCTTCATTTGACACCAAAACCTTCAGAACCTCATGATGCTGTGAG 90
99 gProTyrgIyAlaPheTrpLeuThrAlaAlaPheCysIleLeuSerValL 116
|||||.....|.....|.....|.....|.....|.....|.....|
89 CAGCGCAGGACATTTTGTGATGTCTCGGACGTGTCTTCTTAAATGTGC 40
116 euPheThrLeuThrPheValProGluTrpLysGlyArg 128
|||||.....|.....|.....|.....|.....|.....|.....|
39 TCTTCACGCTTCTTCTTGTCCCGAGACCAAGGCCAA 2

seq_name: gb_est2:BF311341

seq_documentation_block:
LOCUS BF311341 911 bp mRNA EST 21-NOV-2000
DEFINITION 601896741F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4125971 5',
mRNA sequence.
ACCESSION BF311341
VERSION BF311341.1 GI:11259060
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 911)
REFERENCE
  NIH-MGC http://mgi.cni.nih.gov/
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
```



```

1316 TAACAGCATCATGAGATCTCAGACGCTTACGCGCTTCTGCTCACC 1365
      84 laAlapheCysIleuSerValIeuPheTrp 94
      1366 CTGCTTCTGTATCTCTCAGCGCTTCTTTCACG 1397

seq_name: gb_rn:RNO245935

seq_documentation_block:
LOCUS      RNO245935      2087 bp      mRNA      18-FEB-2000
DEFINITION Rattus norvegicus mRNA for glucose transporter (GLUTX1 gene).
ACCESSION  AJ245935
VERSION     AJ245935.1 GI:7018604
KEYWORDS   glucose transporter; GLUTX1 gene.
SOURCE      Norway rat.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 2087)
AUTHORS    Ibberson,M., Uldry,M. and Thorens,B.
TITLE      GLUTX1, a novel mammalian glucose transporter expressed in the
            central nervous system and insulin-sensitive tissues
JOURNAL    J. Biol. Chem. 275 (7), 4607-4612 (2000)
MEDLINE    20138191
AUTHORS    Ibberson,M.R.
TITLE      Direct Submission
REFERENCE   Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
            and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
            CH-1005, SWITZERLAND
FEATURES   location/Qualifiers
            source          1..2087
                        /organism="Rattus norvegicus"
                        /db_xref="taxon:10116"
            gene            30..1466
                        /gene="GLUTX1"
            CDS             30..1466
                        /gene="GLUTX1"
                        /codon_start=1
                        /product="glucose transporter"
                        /protein_id="CAB5729.1"
                        /db_xref="GI:7018605"
                        /translation="MSPEDPQETQPLRSPGARPGRRVLFATPAALGPLSPFAL
                        GYSPAPLSLRTPALRLGDTAASWFGAVTGAAGCVLGLDRAGRLSL
                        CTVPFTGFAVITAAADVMMILGRLITGLACGVALVPISEIAYPVAVRGLGSC
                        VOAMVVTGILLAVAGVLEMRMLAVIGCVPTMLLMCMETPRELLTHOYORA
                        MAALRPLMGSEEGMEPPVGAEHOGFOLAMLRPGVHKPLITGICLAVFOOLSGVNAI
                        MFYANTIFEAKRKDSSLASVTGIIIOVLTAAVALIMDRAGRLILALSGVIVFSM
                        SARGVYFKLTQSPNSSHVGLVPIISAEPADVHLGLAMLAVSGMCLPIAGFAVGMP
                        IPMLMSEIFPLHIKGVATGCVLTNMFMAFLVTKFNSIMEILRPYGAWLTAFCI
                        LSVLEFTLTPVETKGRLEQITAHFEGR"
BASE COUNT 355 a      673 c      580 g      479 t
ORIGIN
alignment_scores:
Quality: 508.00      Length: 94
Ratio: 5.404      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-516-493-12 x RNO245935 ..
Align seg 1/1 to: RNO245935 from: 1 to: 2087

1 GIUPROAlaSPvaIHisIeuGlYLeuAlaTrpLeuAlaValGlySerMe 17
|||||
1116 GAGCTGCTGATGATTCACCTGGGCGCTGCTGCTGCTAGCGACAGAT 1165
|||||
17 TCysIeuPheTrpIleuAlaGlyPheAlaValGlyTrpGlyProIleProTriP 34
|||||
1166 GTGCTTCTGTATCTCTCAGCGCTTCTTTCACG 1215
|||||

```

```

34 euIeuMetSerGluIlePheProLeuHisIleuSclYValAlaIlePrgly 50
|||||
1216 TCCCTCATGTCACAGATCTTCCCTCTGACATCAAGGCTGTGCTACCGC 1265
|||||
51 ValCysValIeuThrAsnTrpPheMetAlaPheLeuValThrIleGluIle 67
|||||
1266 GTCTGTGCTCCTCACCACACTGTTTCATGGCGCTTCTGTGACCAAGAGT 1315
|||||
67 eAnSerrIleMetGluIleLeuArgProIleGlyAlaPheTrpLeuThr 84
|||||
1316 TAACAGCATCATGAGATCTCAGACGCTTACGCGCTTCTGCTCACC 1365
|||||
84 laAlapheCysIleuSerValIeuPheThr 94
      1366 CTGCTTCTGTATCTCTCAGCGCTTCTTTCACG 1397

seq_name: gb_rn:AB033418

seq_documentation_block:
LOCUS      AB033418      2189 bp      mRNA      11-APR-2000
DEFINITION Rattus norvegicus glut8 mRNA for glucose transporter 8, complete
            cds.
ACCESSION  AB033418
VERSION     AB033418.1 GI:7592743
KEYWORDS   glucose transporter 8.
SOURCE      Rattus norvegicus CDNA to mRNA, clone_l1b:testis.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 2189)
AUTHORS    Ishibashi,K.
TITLE      Molecular cloning of a new putative glucose transporter
JOURNAL    Published Only in Database (2000) In press
MEDLINE    2 (bases 1 to 2189)
AUTHORS    Ishibashi,K.
TITLE      Direct Submission
REFERENCE   Submitted (08-OCT-1999) to the DDBJ/EMBL/Genbank databases. Kenichi
            Ishibashi, Jichi Medical School, Pharmacology; Minami-Kawachi,
            Kawachi, Tochigi 329-0498, Japan (E-mail:kishbaejichi.ac.jp,
            Tel:81-285-58-7326, Fax:81-285-44-5541)
FEATURES   location/Qualifiers
            source          1..2189
                        /organism="Rattus norvegicus"
                        /db_xref="taxon:10116"
                        /clone_l1b="testis"
            gene            109..1548
                        /gene="glut8"
            CDS             109..1548
                        /gene="glut8"
                        /note="Glut8 has 12 transmembrane segments"
                        /codon_start=1
                        /product="glucose transporter 8"
                        /protein_id="BA94383.1"
                        /db_xref="GI:7592744"
                        /translation="MSPEDPQETQPLRSPGARPGRRVLFATPAALGPLSPFAL
                        GYSPAPLSLRTPALRLGDTAASWFGAVTGAAGCVLGLDRAGRLSL
                        LCTVPFTGFAVITAAADVMMILGRLITGLACGVALVPISEIAYPVAVRGLGSC
                        VOAMVVTGILLAVAGVLEMRMLAVIGCVPTMLLMCMETPRELLTHOYORA
                        MAALRPLMGSEEGMEPPVGAEHOGFOLAMLRPGVHKPLITGICLAVFOOLSGVNAI
                        MFYANTIFEAKRKDSSLASVTGIIIOVLTAAVALIMDRAGRLILALSGVIVFSM
                        SARGVYFKLTQSPNSSHVGLVPIISAEPADVHLGLAMLAVSGMCLPIAGFAVGMP
                        IPMLMSEIFPLHIKGVATGCVLTNMFMAFLVTKFNSIMEILRPYGAWLTAFCI
                        LSVLEFTLTPVETKGRLEQITAHFEGR"
BASE COUNT 392 a      702 c      609 g      486 t
ORIGIN
alignment_scores:
Quality: 508.00      Length: 94
Ratio: 5.404      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 100.000

```

alignment_block:

us-09-516-493-12 x AB033418 ..

Align seg 1/1 to: AB033418 from: 1 to: 2189

```
1  GIUProAlaAspValHisLeuGlyLeuAlaTrpLeuAlaValAlaGlySerMet 17
   |||
1198 GAGCGTGTGATGTTCCACCTGGGCTGGCTGGCTGTAGGACAGCAT 1247
   |||
17  tCysLeuPheIleAlaGlyPheAlaValAlaGlyTrpGlyProIleProTrpL 34
   |||
1248 GTCCCTCTTCATCGCTGGTTTGGCAGTAGGCTGGGACCCATCCCTGGCC 1297
   |||
34  euleMetSerGluIlePheProLeuHisIleLysGlyValAlaTrpGly 50
   |||
1298 TCCGTCATGTCAGAGATCTTCCCTGTGCACATCAAGCGTGGCTACCGGC 1347
   |||
51  ValCysValLeuThrAsnTrpPheMetAlaPheLeuValThrLysGluP 67
   |||
1348 GTCTGTGTCTCCACCACTGGTTTCATGCTTGTGTCAGCAAGAGTT 1397
   |||
67  eAsnSerIleMetGluIleLeuAlaGProGlyAlaPheTrpLeuThrA 84
   |||
1398 TAACAGCATCATGAGATCTTCAGACCTACGCGCTTGTGGCTACCG 1447
   |||
84  IaAlaPheCysIleLeuSerValLeuPheThr 94
   |||
1448 CTGCTTCTGTATCCTCAAGCTCTTTTCACG 1479
   |||
```

seq_name: gb_pat:AX076671

seq_documentation_block:

LOCUS AX076671 2072 bp DNA PAT 06-FEB-2001
DEFINITION Sequence 5 from Patent WO0104145.
ACCESSION AX076671
VERSION AX076671.1 GI:12711202
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2072)
AUTHORS Thorens,B., Ibberson,M. and Uldry,M.
TITLE Gluc polypeptide family and nucleic acids encoding same
JOURNAL Patent: WO 0104145-A 5 18-JAN-2001;
University of Lausanne (CH)

FEATURES

source

1..2072
/organism="Mus sp."
/db_xref="taxon:10095"
21..1454
/note="unnamed protein product"

CDS

/codon_start=1
/protein_id="CAC28497.1"
/db_xref="GI:12711203"
/translation="MSPEDPQETQPLRPPEARTPRGRVFLASFAALGRLNFGFAL
GSSPAIPSLRTPALRLGDNAASWFGAVVTLGAAAGILGWLDRSGRKSLLL
CTPVTGFAVITTAARDVMLLGGRLTLGACVASIAPVYISEIAYPAVRGLGSC
VOLMVTGILLATVAGCVLEMRMLAVLGCVPPLMLLKCIMPETPRFLTHOYOA
MAALRFLMGSEBGEPPVGAEHQGFOLALRRPGIKPLIGISLWVFOQLSGVNAI
MFYANSIFBEAKFRDSSLASVTGIIQVLTAAVALMDRAGRRLILALSGLVIMVFSM
SAFGTYFKLQSLPSNSSHVGIVPIAAEPYDVOVGILAVGSMCLFIAGFAVGWGP
PMLMSEIFPLHKGATGICVLTNNMFALFVTKENSVEMELRPYGAPMLTAAFCAL
SVLFTLVVPETKGRLEOVTAHEEGR"

BASE COUNT 351 a 673 c 584 g 464 t
ORIGIN

alignment_scores:

Quality: 482.00 Length: 94
Ratio: 5.183 Gaps: 0
Percent Similarity: 98.936 Percent Identity: 91.489

alignment_block:

us-09-516-493-12 x AX076671 ..

Align seg 1/1 to: AX076671 from: 1 to: 2072

```
1  GIUProAlaAspValHisLeuGlyLeuAlaTrpLeuAlaValAlaGlySerMet 17
   |||
1104 GAGCGTGTGATGTTCCACCTGGGCTGGCTGGCTGTAGGACAGCAT 1153
   |||
17  tCysLeuPheIleAlaGlyPheAlaValAlaGlyTrpGlyProIleProTrpL 34
   |||
1154 GTCCCTCTTCATCGCTGGTTTGGCAGTAGGCTGGGACCCATCCCTGGCC 1203
   |||
34  euleMetSerGluIlePheProLeuHisIleLysGlyValAlaTrpGly 50
   |||
1204 TCCGTCATGTCAGAGATCTTCCCTGTGCACATCAAGCGTGGCTACCGGC 1253
   |||
51  ValCysValLeuThrAsnTrpPheMetAlaPheLeuValThrLysGluP 67
   |||
1254 ATCTGTGTCTCCACCACTGGTTTCATGCTTGTGTCAGCAAGAGTT 1303
   |||
67  eAsnSerIleMetGluIleLeuAlaGProGlyAlaPheTrpLeuThrA 84
   |||
1304 CAACACCGTCATGAGATCTTCAGACCTACGCGCTTGTGGCTACCG 1353
   |||
84  IaAlaPheCysIleLeuSerValLeuPheThr 94
   |||
1354 CTGCTTCTGTATCCTCAAGTGTCTTATTCACA 1385
   |||
```

seq_name: gb_ro:MMU245936

seq_documentation_block:

LOCUS MMU245936 2072 bp mRNA ROD 18-FEB-2000
DEFINITION Mus musculus mRNA for glucose transporter (GLUTX1 gene).
ACCESSION AJ245936
VERSION AJ245936.1 GI:7018606
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2072)
AUTHORS Ibberson,M., Uldry,M. and Thorens,B.
TITLE GLUTX1, a novel mammalian glucose transporter expressed in the
central nervous system and insulin-sensitive tissues
JOURNAL J. Biol. Chem. 275 (7), 4607-4612 (2000)

FEATURES

source

1..2072
/organism="Mus musculus"
/db_xref="taxon:10090"
21..1454
/gene="GLUTX1"
21..1454
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CA875719.1"
/db_xref="GI:7018607"

CDS

/translation="MSPEDPQETQPLRPPEARTPRGRVFLASFAALGRLNFGFAL
GSSPAIPSLRTPALRLGDNAASWFGAVVTLGAAAGILGWLDRSGRKSLLL
CTPVTGFAVITTAARDVMLLGGRLTLGACVASIAPVYISEIAYPAVRGLGSC
VOLMVTGILLATVAGCVLEMRMLAVLGCVPPLMLLKCIMPETPRFLTHOYOA
MAALRFLMGSEBGEPPVGAEHQGFOLALRRPGIKPLIGISLWVFOQLSGVNAI
MFYANSIFBEAKFRDSSLASVTGIIQVLTAAVALMDRAGRRLILALSGLVIMVFSM
SAFGTYFKLQSLPSNSSHVGIVPIAAEPYDVOVGILAVGSMCLFIAGFAVGWGP
PMLMSEIFPLHKGATGICVLTNNMFALFVTKENSVEMELRPYGAPMLTAAFCAL
SVLFTLVVPETKGRLEOVTAHEEGR"

JOURNAL Submitted (07-FEB-2000) OB/Gyn, Washington University, 4566 Scott Ave, St. Louis, MO 63110, USA

FEATURES

source
1.1843
/organism="Mus musculus"
/strain="129"
/db_xref="taxon:10090"
1.48
49.1482
/codon_start=1
/product="glucose transporter GLUT8"
/protein_id="AAE78366.1"
/db_xref="GI:8671758"
/translation="MSPEDPQETQPLRPPEARPPRRVFLASFAALGPLSPFGL
GSSPAIPSLRTPAPALRLGDNAAMWFGAVVTLGAAGLGLGMLIDRSGRLLSL
CYPPVTFGAVITARDVWMLGRLITGLAGCVASLAVPVYISETAPVRLGLSC
VOLAVYTGILATVAGVLEWRMLAVIGCYPTLMLLMCTMEPTPRELLTQHYDA
MALRLWMSSEGEWEPVGAHQFOLALRRGYIKPLITGSLWFOQLSVNAI
MFVANSIFEEAKRDSLSAVTVGIQVLTAAVALIMDRAGRLLALAGVIMVPS
SAGTYFKLTQSLPSNSHVLPIAAEPDVQVLAIVGSKCLFIAGFAVGMGP
PWLMSIEIPLHVGVATGICVLTNMFMAFLVTKKEFSVEMLRPGAFMLTAACAL
SVLFTLVVPEPTGRTLEQVTAHFEGR"
1483..1843

5'UTR
CDS

3'UTR
BASE COUNT 291 a 609 c 529 g 410 t 4 others
ORIGIN

alignment_scores:

Quality: 477.00 Length: 94
Ratio: 5.129 Gaps: 0
Percent Similarity: 98.936 Percent Identity: 90.426

alignment_block:

US-09-516-493-12 x AF232061 ..

Align seg 1/1 to: AF232061 from: 1 to: 1843

```
1 GIUProAlaSPValHisLeuGlyLeuAlaTrpLeuAlaValGlySerMet 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1132 GAGCCTGTGATGTCACAGTGAGGACTGGCTGGCTGTGAGCAGCAT 1181
17 TCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpL 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1182 GTGCTCTTCATGCTGGCTTTGGCGTGGCGGACCCATCCCTGGCC 1231
34 euleuMetSerGluIlePheProLeuHisIleGlyValAlaIleThrGly 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1232 TCCTATATGTCAGAGATCTTCCCTCTGCATGTCAAGGGTGGCTACCGC 1281
51 ValCysValLeuThrAsnTrpPheMetAlaPheLeuValThrIleGlyLuh 67
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1282 ATCTGTGTCACCACTGCTTATGCGCTTCTAGTGACCAAGAGATT 1331
67 eanSerIleMetGluIleLeuArgProTrpGlyAlaPheTrpLeuThrA 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1332 CAGCAGCTCATGTGAGATGCTCAGACCTACGCTTCTGCTACACGC 1381
84 IaAlaPheCysIleLeuSerValLeuPheThr 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1382 CTGCTTCTGGCGCTCTCAGTGTCTATTACACA 1413
```

seq_name: gb_om:AF321324

seq_documentation_block:

LOCUS AF321324 1012 bp mRNA MAM 02-JUL-2001
DEFINITION Bos taurus glucose transporter 8 (Glut8) mRNA, partial cds.
ACCESSION AF321324
VERSION AF321324.1 GI:14582715
KEYWORDS
SOURCE
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;

REFERENCE
1 (bases 1 to 1012)
AUTHORS Augustin, R., Navarrete-Santos, A. and Fischer, B.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2000) Anatomy and Cell Biology, Faculty of
Medicine, Martin Luther University Halle, Grosse Steinstrasse 52,
Halle 061097, Germany

FEATURES

source
1.1012
Location/Qualifiers
/organism="Bos taurus"
/db_xref="taxon:9913"
<1..1012
/gene="Glut8"
<1..1005
CDS
/gene="Glut8"
/codon_start=1
/product="glucose transporter 8"
/protein_id="AAK69606.1"
/db_xref="GI:14582716"
/translation="ITSEIAYPEVNGILGSCVOLAVYTGILATVAGVLEWRMLAVLG
CVAPSPMLLMCEPMPTRPRLISQKHQDEAMAOPLMGVAQGEWEPPLGQODPFV
AQLRRPGVYKPRITIGSLMAFOQLSGVNAVYFAETLFEPAKFKDSSLSAVNGVIOY
LFTATNALIMDRAGRLLTLDSGVVWVFSTISACTGYKILGEGPSNSHDLPLVSM
EADPTNGLAMLAIGNMCLFAGFAVGMGPWLMSIEIPLHVGVATGCVLTNMF
MAFLVTKKEFSLSLMEVLRPGAFWLASAFICFVGLFTACVPETKGKTLQITAHFEGR"

BASE COUNT 153 a 323 c 302 g 234 t
ORIGIN

alignment_scores:

Quality: 458.00 Length: 94
Ratio: 4.925 Gaps: 0
Percent Similarity: 98.936 Percent Identity: 86.170

alignment_block:

US-09-516-493-12 x AF321324 ..

Align seg 1/1 to: AF321324 from: 1 to: 1012

```
1 GIUProAlaSPValHisLeuGlyLeuAlaTrpLeuAlaValGlySerMet 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
655 GAGCGCGCGGATACCAACGTGGGGCTGGCTGGCGTGGCGGCAACAT 704
17 TCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpL 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
705 GTGCTCTTCATGCGCGCTTGGCTGGCGTGGGGCCCATCCCTGGCG 754
34 euleuMetSerGluIlePheProLeuHisIleGlyValAlaIleThrGly 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
755 TCCTATATGTCAGAGATCTTCCCTCTGCATGTCAAGGGTGGCTACCGC 804
51 ValCysValLeuThrAsnTrpPheMetAlaPheLeuValThrIleGlyLuh 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
805 GTCTGCTGTCACCAACGTGTCATGCGCTTCTGTCACCAAGAGATT 854
67 eanSerIleMetGluIleLeuArgProTrpGlyAlaPheTrpLeuThrA 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
855 CAGCACTCATGTGAGATGCTCAGGCCCTTACGGTGGCTTCTGGCTGGCT 904
84 IaAlaPheCysIleLeuSerValLeuPheThr 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
905 CTGCTTCTGGCGCTCTCAGTGTCTATTACACT 936
```

seq_name: gb_pr:HS17801

seq_documentation_block:

LOCUS HS17801 1508 bp mRNA PRI 13-MAY-2000
DEFINITION Homo sapiens mRNA for glucose transporter 8 (Glut8 gene).
ACCESSION Y17801
VERSION Y17801.1 GI:7688145
KEYWORDS glucose transporter 8; GLUT8 gene.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1508)
AUTHORS Doege, H., Schumann, A., Bahrenberg, G., Brauers, A. and Joost, H.G.
TITLE GLUT8, a novel member of the sugar transport facilitator family
JOURNAL with glucose transport activity
MEDLINE 20283667
REFERENCE 2 (bases 1 to 1508)
AUTHORS Joost, H.G.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1998) H. Joost, Institute of Pharmacology and
Toxicology, Technical University Aachen, Wendlingweg 2, D-52057
Aachen, FRG

FEATURES
source Location/Qualifiers
1..1508
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue-type="testis"
27..1460
/gene="GLUT8"
27..1460
/gene="GLUT8"
/codon_start=1
/product="glucose transporter 8"
/protein_id="CAB89809.1"
/db_xref="GI:7688146"
/translation="MPEDPPEPILGREGSAPRGKRVFLAFAALSPISFGL
CYSPALPISIORAPAPRILDDAASWFGAVVTIGAAAGVGLGMLVDKGRKLSLL
CSVPVAGFAYITAGQVWMLLGRLLTGLAGVASIAPVYISEIAPVAVGLGSC
VOLWVVGILAYLAGVWLEWRMLVGLCVPSLMLLFCMPETPRILLRORLEA
MAALRFLEAKFSDGDDPPIGAESFHLALRQPSLYKPEIIGVSLMFQOLSGVNAV
FYATTFIEEAKFKDSSLASVYGVIOVFLFTVALINDRGRLLIVLISGVVWFST
AFGAYFKLTGGPGNSHVASAPNSAPVAGSLAMLAGSMCLFTAGAVMGPI
PMLMSEIFPLHVGAVATGICVLTNMLIAFLVLTKEFSLMELRPGAFWLASFCIF
SVLEFSCVPRTKGLQITAHFEGR"

BASE COUNT 194 a 519 c 477 g 318 t
ORIGIN

alignment_scores:
Quality: 457.00 Length: 94
Ratio: 4.914 Gaps: 0
Percent Similarity: 98.936 Percent Identity: 85.106

alignment_block:
US-09-516-493-12 x HSA17801 ..

Align seg 1/1 to: HSA17801 from: 1 to: 1508

```

1 GIUPTOLAASPVAlHIsleuGlyLeuAlATrPLeuAlaValGlySerMe 17
  ::::::::::::::: :::::::::::::::
1110 CAGCCTGTGATGACCAAGCGTGGGCGCTGGCGTGGCGGACGAT 1159
17 tCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpL 34
  ::::::::::::::: :::::::::::::::
1160 GGGCCCTTCATGCGCGGCTTGGGCTGGGCGGCCCATGCCCTGGC 1209
34 euLeuMeSerGluIlePheProLeuHisIleLysGlyValAlaThrGly 50
  ::::::::::::::: :::::::::::::::
1210 TCCTCATGTCAGAGATCTTCCTGTCATGCAAGGCGTGGGACAGGC 1259
51 ValCysValLeuThrAsnTrpPheMetAlaPheLeuValThrLysGluPh 67
  ::::::::::::::: :::::::::::::::
1260 ATCTGCGTCTTACCACTGCTCATGGCTTCTCTGACCAAGGAGATT 1309
67 eaSnSerIleMeTgLuIleLeuArgProTyrGlyAlaPheTrpLeuThra 84
  ::::::::::::::: :::::::::::::::
1310 CAGCAGCCTCATGAGAGTCTCAGGCGCTATGAGGCTTCTGGCTTGCT 1359
84 laAlaPheCysIleLeuSerValLeuPheThr 94
  ::::::::::::::: :::::::::::::::

```

1360 CCGCTTTTCGACATCTTCAGTGTCTTTTCACCT 1391
seq_name: gb_pat:AX179740

seq_documentation_block:
LOCUS AX179740 2080 bp DNA PAT 06-AUG-2001
DEFINITION Sequence 28 from Patent WO0146258.
ACCESSION AX179740
VERSION AX179740.1 GI:15132104
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2080)
AUTHORS Baughn, M.R., Burford, N., Au-Young, J., Lu, D.A., Yang, J., Reddy, R.,
Lal, P., Hillman, J.L., Azimzal, Y., Yue, H., Nguyen, D.B., Yao, M.G.,
Gandhi, A.R., Tang, Y.T., and Khan, F.A.
TITLE Transporters and ion channels
JOURNAL Patent: WO 0146258-A 28 28-JUN-2001;
Incyte Genomics, Inc. (US)

FEATURES
source Location/Qualifiers
1..2080
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 1416107CBI"

BASE COUNT 300 a 688 c 632 g 460 t
ORIGIN

alignment_scores:
Quality: 457.00 Length: 94
Ratio: 4.914 Gaps: 0
Percent Similarity: 98.936 Percent Identity: 85.106

alignment_block:
US-09-516-493-12 x AX179740 ..

Align seg 1/1 to: AX179740 from: 1 to: 2080

```

1 GIUPTOLAASPVAlHIsleuGlyLeuAlATrPLeuAlaValGlySerMe 17
  ::::::::::::::: :::::::::::::::
1129 CAGCCTGTGATGACCAAGCGTGGGCGCTGGCGTGGCGGACGAT 1178
17 tCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpL 34
  ::::::::::::::: :::::::::::::::
1179 GTCCTCTTCATGCGCGGCTTGGGCTGGGCGGCCCATGCCCTGGC 1228
34 euLeuMeSerGluIlePheProLeuHisIleLysGlyValAlaThrGly 50
  ::::::::::::::: :::::::::::::::
1229 TCCTCATGTCAGAGATCTTCCTGTCATGCAAGGCGTGGGACAGGC 1278
51 ValCysValLeuThrAsnTrpPheMetAlaPheLeuValThrLysGluPh 67
  ::::::::::::::: :::::::::::::::
1279 ATCTGCGTCTTACCACTGCTCATGGCTTCTCTGACCAAGGAGATT 1328
67 eaSnSerIleMeTgLuIleLeuArgProTyrGlyAlaPheTrpLeuThra 84
  ::::::::::::::: :::::::::::::::
1329 CAGCAGCCTCATGAGAGTCTCAGGCGCTATGAGGCTTCTGGCTTGCT 1378
84 laAlaPheCysIleLeuSerValLeuPheThr 94
  ::::::::::::::: :::::::::::::::
1379 CCGCTTTTCGACATCTTCAGTGTCTTTTCACCT 1410

```

seq_name: gb_pr: HSA245937

seq_documentation_block:
LOCUS HSA245937 1873 bp mRNA PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA for glucose transporter (GLUTX1 gene).
ACCESSION AJ245937
VERSION AJ245937.1 GI:7018305
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1873)
AUTHORS Ibberson, M., Uldry, M., and Thorens, B.
TITLE GLUT1, a novel mammalian glucose transporter expressed in the
central nervous system and insulin-sensitive tissues
JOURNAL J. Biol. Chem. 275 (7), 4607-4612 (2000)
MEDLINE 20138191
REFERENCE 2 (bases 1 to 1873)
AUTHORS Ibberson, M. R.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1999) Ibberson M. R., Institute of Pharmacology
and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
CH-1005, SWITZERLAND
FEATURES
source location/Qualifiers
1..1873 /organism="Homo sapiens"
/db_xref="taxon:9606"
4..1437 /gene="GLUTX1"
4..1437 /gene="GLUTX1"
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CAI5702.1"
/db_xref="GI:7018306"
/translation="MPEDPEEPDPLGPPGSAAPRGRRVFLAFAAALGPIISFPA
LGYSSPAIPSLORAPAPAPRLDDAAASWGAAGVITLGAAGVGLGMLVDRGRKLSILL
CSVPYAGFAVITTAADVMMLGGRLLTGACGASIVAPYISEIAYPAVRGLGSC
VOLNVVGLIATLAGVLEWRMLAVGCVPSLMLLMCFMPTPRFLITQHRQEA
MAALRFPLMSGOEDPPIGAESFHLALRQPIYKPTIGVSLMAFOQLSGVNAV
FYAETTFEAKFKDSSLASVVGIVQLFTFAVALIMDRAGRRLILVSGVAVFST
AFGAYKFIQGGPNSSHVAISAPVASOPDASVGLMALVAGNMLCFIAGPAGVGMPI
PMLMSFPLPHVKVATGICVLNMLMAFLVTRKESSLMEVLRPGAFMLASAFCLF
SVLFTLFCVPEIKKTLQITAHREGH"
BASE COUNT 279 a 621 c 573 g 400 t
ORIGIN
alignment_scores:
Quality: 454.00 Length: 94
Ratio: 4.882 Gaps: 0
Percent Similarity: 98.936 Percent Identity: 84.043
alignment_block:
US-09-516-493-12 x HSA245937 ..
Align seg 1/1 to: HSA245937 from: 1 to: 1873
1 GIUPROALASPVALHISLEUGLYLEUALATPRLEUALAVALGYSERME 17
:::|||||::: |||
1087 CAGCGTGTGATGCCAGCGGTGGCGCTGGCTGGCGCGCTGGCAACT 1136
17 TCYSLEUPHEILAEALAGLYPHEALAVAGLYTRPGLYPROLLEPROTPI 34
|||||
1137 GTGCCTTTCATCGCGCGCTTGGCGGTGGCGGCCCATCCCTGGC 1186
34 EULEUMETSERGLUILEPHERPROLEHISILEYSLYVALAIAATHRGY 50
|||||
1187 TCCCTATGTCACAGATCTTCCCTGCGATGTCAGGCGCTGGCGACAGC 1236
51 VALCYSAVLEUTHRASNTPRPHEMETALAPHELEUVALTHLYSGLIUPH 67
:::|||||::: |||
1237 ATCTGCGTCTACCACTGGCTGATGGCTTTCTCGTACCAAGAGATT 1286
67 EASNSERTLEMETGLUILEUARGPROTYRGYIALAPHETRPLEUTHRA 84
:::|||||::: |||
1287 CAGCAGCCTCATGAGGTCTCTCAGGCCCTATGAGCCCTTCTGCTGGCT 1336
84 LAALAPHCYSTILEUSERVALLEUPHETHR 94
:::|||||::: |||
1337 CCGCTTCTGCATCTTCAGTGTCTTTGCACT 1368

seq_name: gb_pat:AX076667
seq_documentation_block:
LOCUS AX076667 2217 bp DNA PAT 06-FEB-2001
DEFINITION Sequence 1 from Patent WO0104145.
ACCESSION AX076667
VERSION AX076667.1 GI:12711198
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 2217)
AUTHORS Thorens, B., Ibberson, M., and Uldry, M.
TITLE Glutx polypeptide family and nucleic acids encoding same
JOURNAL Patent: WO 0104145-A 1 18-JAN-2001;
University of Lausanne (CH)
FEATURES
source location/Qualifiers
1..2217 /organism="Homo sapiens"
/db_xref="taxon:9606"
348..1781 /note="unnamed protein product"
/codon_start=1
/protein_id="GI:12711199"
/db_xref="GI:12711199"
/translation="MPEDPEEPDPLGPPGSAAPRGRRVFLAFAAALGPIISFPA
LGYSSPAIPSLORAPAPAPRLDDAAASWGAAGVITLGAAGVGLGMLVDRGRKLSILL
CSVPYAGFAVITTAADVMMLGGRLLTGACGASIVAPYISEIAYPAVRGLGSC
VOLNVVGLIATLAGVLEWRMLAVGCVPSLMLLMCFMPTPRFLITQHRQEA
MAALRFPLMSGOEDPPIGAESFHLALRQPIYKPTIGVSLMAFOQLSGVNAV
FYAETTFEAKFKDSSLASVVGIVQLFTFAVALIMDRAGRRLILVSGVAVFST
AFGAYKFIQGGPNSSHVAISAPVASOPDASVGLMALVAGNMLCFIAGPAGVGMPI
PMLMSFPLPHVKVATGICVLNMLMAFLVTRKESSLMEVLRPGAFMLASAFCLF
SVLFTLFCVPEIKKTLQITAHREGH"
BASE COUNT 320 a 750 c 672 g 475 t
ORIGIN
alignment_scores:
Quality: 454.00 Length: 94
Ratio: 4.882 Gaps: 0
Percent Similarity: 98.936 Percent Identity: 84.043
alignment_block:
US-09-516-493-12 x AX076667 ..
Align seg 1/1 to: AX076667 from: 1 to: 2217
1 GIUPROALASPVALHISLEUGLYLEUALATPRLEUALAVALGYSERME 17
:::|||||::: |||
1431 CAGCGTGTGATGCCAGCGGTGGCGCTGGCTGGCGCGCTGGCAACT 1480
17 TCYSLEUPHEILAEALAGLYPHEALAVAGLYTRPGLYPROLLEPROTPI 34
|||||
1481 GTGCCTTTCATCGCGCGCTTGGCGGTGGCGGCCCATCCCTGGC 1530
34 EULEUMETSERGLUILEPHERPROLEHISILEYSLYVALAIAATHRGY 50
|||||
1531 TCCCTATGTCACAGATCTTCCCTGCGATGTCAGGCGCTGGCGACAGC 1580
51 VALCYSAVLEUTHRASNTPRPHEMETALAPHELEUVALTHLYSGLIUPH 67
:::|||||::: |||
1581 ATCTGCGTCTACCACTGGCTGATGGCTTTCTCGTACCAAGAGATT 1630
67 EASNSERTLEMETGLUILEUARGPROTYRGYIALAPHETRPLEUTHRA 84
:::|||||::: |||
1631 CAGCAGCCTCATGAGGTCTCTCAGGCCCTATGAGCCCTTCTGCTGGCT 1680
84 LAALAPHCYSTILEUSERVALLEUPHETHR 94
:::|||||::: |||
1681 CCGCTTCTGCATCTTCAGTGTCTTTGCACT 1712

```

seq_name: gb_pr:AL445222
seq_documentation_block:
LOCUS      AL445222      225370 bp      DNA      PRI      24-Apr-2001
DEFINITION Human DNA sequence from clone RP11-356B19 on chromosome 9, complete
sequence.
ACCESSION  AL445222
VERSION     AL445222.9
KEYWORDS    HG.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE   1 (bases 1 to 225370)
AUTHORS     Laird, G.
TITLE       Direct Submission
JOURNAL     Submitted (24-Apr-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
            requests: clonerequests@sanger.ac.uk
            On Apr 26, 2001 this sequence version replaced gi:13277497.
COMMENT     During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em, EMBL; Sw,
            SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
            Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr9
            RP11-356B19 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBac3.6
            This sequence is the entire insert of clone RP11-356B19. The true
            left end of clone RP11-373J8 is at 96439 in this sequence. The true
            right end of clone RP13-225O21 is at 9980 in this sequence.
FEATURES             Location/Qualifiers
     source           1..225370
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="9"
                     /clone="RP11-356B19"
     misc_feature     23455..23821
                     /clone_id="RPCI-11.2"
                     /note="Single clone region. Assembly confirmed by
                     restriction digest data."
     misc_feature     65019..65283
                     /note="Single clone region. Sequence from reads from a
                     short insert library derived from a single pUC clone.
                     Restriction digest data confirm the assembly."
     misc_feature     225123..225370
                     /note="Single clone region. Sequence from reads from a
                     short insert library derived from a single pUC clone.
                     Restriction digest data confirm the assembly."
BASE COUNT          56329 a 57092 c 57267 g 54682 t
ORIGIN
alignment_scores:

```

```

Quality: 266.50      Length: 70
Ratio: 4.675
Percent Similarity: 81.429      Percent Identity: 72.857

alignment_block:
US-09-516-493-12 x AL445222      ..
Align seg 1/1 to: AL445222 from: 1 to: 225370

2 ProAlaAspValHisLeuGlyLeuAlaTrpLeuAlaValGlySerMetCys 18
|||||
61171 CCGCGGACCTGATGCTGCTGCGC.....TG 61199
18 sLeuPheIleAlaGlyPheAlaValAlaGlyTrpGlyProIleProTrpLeu 35
|||||
61200 TCTCGCTCCACAGAGGCTTTCGGGTGGCGTGGGGCCCATCCCTGCTCC 61249
35 eMetSerGluIlePheProLeuHisIleLysGlyValAlaThrGlyVal 51
|||||
61250 TCATGTCAAGAGATCTCCCTGTCATGTCACAGGCGCTGGCGACAGGCATC 61299
52 CysValLeuThrAsnTrpPheMetAlaPheLeuValThrLysGluPheAs 68
|||||
61300 TCGGCTCTACCACTGAGGCTCATGGCTTCTCGTGAACCAAGAGTTTCAG 61349
68 nSerIleMet 71
|||||
61350 CAGCTCATG 61359

seq_name: gb_pat:AX072418
seq_documentation_block:
LOCUS      AX072418      378 bp      DNA      PAT      25-JAN-2001
DEFINITION Sequence 2890 from Patent WO0102568.
ACCESSION  AX072418
VERSION     AX072418.1
KEYWORDS    AX072418.1 GI:12582769
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE   1 (bases 1 to 378)
AUTHORS     Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.,
            Kassam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D.,
            Lamsom, G., Drmanac, R., Crkenjakov, R., Drmanac, S., Dickson, M.,
            Labat, I., Ieshkowitz, D., Kita, D., Garcia, Y. and Strache-Crain, B.
            Human genes and gene expression products
            Patent: WO 0102568-A 2890 11-JAN-2001;
            CHIRON CORPORATION (US); HYSPO, INC. (US)
FEATURES             Location/Qualifiers
     source           1..378
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT          45 a 130 c 117 g 86 t
ORIGIN
alignment_scores:
Quality: 221.00      Length: 77
Ratio: 3.348
Percent Similarity: 85.714      Percent Identity: 59.740

alignment_block:
US-09-516-493-12 x AX072418      ..
Align seg 1/1 to: AX072418 from: 1 to: 378

19 LeuPheIleAlaGlyPheAlaValAlaGlyTrpGlyProIleProTrpLeu 35
|||||
31 CTCCTCATCATGCGCTACGCGGCTGGCGTGGCGCCATCATCCTGCTGCT 80
35 uMetSerGluIlePheProLeuHisIleLysGlyValAlaThrGlyValC 52
|||||

```


81	CATGCTGAGAGTCTCCCTCGCTCCCGCTGCGCTCAGGGCTCT	130
52	y\$ValLeuThr\$nr\$trp\$hem\$e\$al\$aph\$e\$u\$Val\$thr\$lys\$g\$lu\$phe\$asn	68
131	GGCTGCTGGCCAGCTGCGTACCAGCTTGCTCCTACCAAGTCTTCG	180
69	Se\$1\$le\$Me\$Gl\$u\$le\$u\$Ar\$g\$Pro\$Tyr\$Gly\$Ala\$Phe\$Trp\$leu\$Thr	Ala 85
181	CCAGTCTGCT GAGCACCCTTCGCGCTCCAGCGCTTCCTTCTCTTCGCGG	229
85	l\$aph\$e\$st\$le\$u\$e\$u\$er\$Val\$leu\$Phe\$Thr	94
230	CCATCTGCTTGGAGCTGCTGCTTCACA	258
seq_name:	qb_pr: HSA17803	
seq_documentation_block:	1580 bp	mRNA
LOCUS	HS17803	PR1
DEFINITION	Homo sapiens mRNA for facilitative glucose transporter 6, GLUT6 (SLC2A6 gene).	13-JUL-2001
ACCESSION	Y17803	
KEYWORDS	Y17803.1 GI:9187481	
SOURCE	glucose transporter 6, GLUT6; SLC2A6 gene.	
ORGANISM	human.	
REFERENCE	Homo sapiens	
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE	1 (bases 1 to 1580)	
JOURNAL	Dodge, H., Boclanski, A., Joost, H.G. and Schurrmann, A.	
MEDLINE	Activity and genomic organization of human glucose transporter 9 (GLUT9), a novel member of the family of sugar-transport facilitators predominantly expressed in brain and leucocytes	
AUTHORS	Biochem. J. 350 Pt 3, 771-776 (2000)	
REFERENCE	20427701	
JOURNAL	2 (bases 1 to 1580)	
TITLE	Joost, H.G.	
AUTHORS	Direct Submission	
REFERENCE	Submitted (16-JUL-1998) H. Joost, Institute of Pharmacology and Toxicology, Technical University Aachen, Wendingweg 2, D-52057 Aachen, FRG	
FEATURES	Location/Qualifiers	
Source	1..1580	
gene	/organism="Homo sapiens"	
CDS	/db_xref="taxon:9606"	
	/chromosome="9"	
	/map="9q34"	
	/cell_type="leukocytes"	
	47..1570	
	/gene="SLC2A6"	
	47..1570	
	/gene="SLC2A6"	
	/codon_start=1	
	/product="Facilitative glucose transporter 9 (GLUT6)"	
	/protein_id="CAB96996.1"	
	/db_xref="GI:9187482"	
	/db_xref="SPRMBL:O9UCQ3"	
	/translation="MOEPLGCAEGPDVDTPEKPPSPGDRARVGLTQNRKPLATFAVIGNFSFGVALVYTSVIPALERSLDPLRLTSQASMFSGSVFTGCAAGCSAMITINDLGRKLSIMFSAVPSPAGCALMAGAHGIMLLIGRTITGFACTLRACIPVYSEAPPCVRGALGATPOLMAVVFSGLSLYALGLLPMWLVAVGAPVLMILITPMNSBREFLSGRDEDEBALRALMLRGTDDVHMEFQIDQNVKQSSVMAEARAPHVCRRPITVALIMRLQLOLTGITPTILVYLOSIFPSTDAIIVGAVRLSTVIALTMDMLAGKRWLLPVSAAIMFPAANLTGLYIHGPPRLSPNSAGLESQSWDLAOPAPAGVYTLITPLATITMLPTMGYAVAGMGPTIMLLMSVPLPARGVASGICVLASMTATVLRKSFLEPVYSTFGIVQYFFPFPFPAALCIYSLVPTGCGVETFGKRSLEQIESFFRMGRSFLR"	
BASE COUNT	216 a 552 c 492 g 320 t	
ORIGIN		
alignment_scores:		
Quality:	221.00	Length: 77
Ratio:	3.348	Gaps: 1
Percent Similarity:	85.714	Percent Identity: 59.740

[illegible]

Thu Feb 14 07:44:23 2002

us-09-516-493-12.rge

Page 10

17 tCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpL 34
 1166 GGGCTTTCATCGCGTGTGGAGTAGGCTGGGACCATCCCTGGC 1215
 34 euLeuMetSerGluIlePheProLeuHisIleLysGlyAlaIleTrpGly 50
 1216 TCCTCATGTCAGAGATCTTCCTCCACATCAGAGGTGGTGCTACCGC 1265
 51 ValGlyValLeuThrAsnTrpPheMetAlaPheLeuValThrLysGly 67
 1266 GTCTGTGCTCCTACCACTGGTTCATGGCTTTCGTGACCAAGAGTT 1315
 67 eaSnSerIleMetGluIleLeuArgProTyrGlyAlaPheTrpLeuThra 84
 1316 TACAGGATCATGAGATCTTCAGACCCCTACGGCCCTTCTGGCTCACCG 1365
 84 IaAlaPheCysIleLeuSerValLeuPheThr 94
 1366 CTGCCTTCGTATCCTCAGCGCTCTTTCACG 1397
 seq_name: /SIDS2/gcdata/geneseq/NA2001.DAT:AAF55867

seq_documentation_block:
 ID AAF55867 standard; cDNA: 2072 BP.
 AC AAF55867;
 DT 17-APR-2001 (first entry)
 DE Murine GLUTX1 coding sequence.
 KW Murine; GLUTX: gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 OS Mus sp.
 PN WO200104145-A2.
 PD 18-JAN-2001.
 PF 14-JUL-2000; 2000WO-IB01042.
 PR 14-JUL-1999; 99US-0143907.
 PR 27-AUG-1999; 99US-0151140.
 PR 23-FEB-2000; 2000US-0184285.
 PR 13-JUL-2000; 2000US-0616132.
 PA (UCLA-) UNIV LAUSANNE.
 PI Thorens B, Ibberson M, Udry M;
 DR WPI; 2001-112615/12.
 DR P-PSDB; AAB66934.
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischaemia and diabetes -
 PS Claim 3; Page 73-74; 124pp; English.
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is the coding sequence
 CC for murine GLUTX1.
 XX Sequence 2072 BP; 351 A; 673 C; 584 G; 464 T; 0 other;
 SQ

alignment_scores:
 Quality: 482.00 Length: 94
 Ratio: 5.183 Gaps: 0
 Percent Similarity: 98.936 Percent Identity: 91.489

alignment_block:
 US-09-516-493-12 x AAF55867 ..

Align seg 1/1 to: AAF55867 from: 1 to: 2072

1 GluProAlaAspValHisLeuGlyLeuAlaTrpLeuAlaValGlySerLe 17
 1104 GAACCTGTGGATGTCACAGTGGAGTGGCTGGCTGGCTGTAGGACAT 1153
 17 tCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpL 34
 1154 GTGCTTTCATGCTGCTGGCTTTCGGGTGGCTGGGACCATCCCTGGC 1203
 34 euLeuMetSerGluIlePheProLeuHisIleLysGlyAlaIleTrpGly 50
 1204 TCCTCATGTCAGAGATCTTCCTCCATGTCAGAGGTGGTGCTACCGC 1253
 51 ValGlyValLeuThrAsnTrpPheMetAlaPheLeuValThrLysGly 67
 1254 ATCTGTGCTCCTACCACTGGTTCATGGCTTTCGTAGTACCAAGAGTT 1303
 67 eaSnSerIleMetGluIleLeuArgProTyrGlyAlaPheTrpLeuThra 84
 1304 CAACAGCGTATGAGATGCTCAGACCCCTACGGTGGCTTTCGCTCACCG 1353
 84 IaAlaPheCysIleLeuSerValLeuPheThr 94
 1354 CTGCCTTCGCGCTCAGTGTCTTATTCACA 1385
 seq_name: /SIDS2/gcdata/geneseq/NA2001.DAT:AAD09552

seq_documentation_block:
 ID AAD09552 standard; cDNA: 2080 BP.
 AC AAD09552;
 DT 10-SEP-2001 (first entry)
 DE Human transporter and ion channel-1 (TRICH-1) cDNA.
 DE XX
 KW Human; transporter and ion channel-1; TRICH-1; vaccine; cystic fibrosis;
 KW gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy;
 KW hypertension; angina; neurological disorder; asthma; bipolar disorder;
 KW dementia; depression; Alzheimer's disease; epilepsy; mood; arthritis;
 KW Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke;
 KW Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy;
 KW demyelinating disease; mental disorder; schizophrenia; polymyositis;
 KW muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease;
 KW dermatomyositis; diabetes mellitus; immunological disorder; psoriasis;
 KW rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus;
 KW sickle cell anaemia; Wilson's disease; infertility; Cushing's disease;
 KW scleroderma; pulmonary artery stenosis; noctropic; Addison's disease;
 KW malabsorption syndrome; hypercholesterolaemia; cancer; ss.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH CDS 46..1479
 FT /*tag= a
 FT /product= "Human TRICH-1 protein"
 PN WO200146258-A2.
 PD 28-JUN-2001.
 XX 22-DEC-2000; 2000WO-US35095.
 PF 23-DEC-1999; 99US-0172000.
 XX

```

PR 14-JAN-2000: 2000US-0176083.
PR 21-JAN-2000: 2000US-0177332.
PR 28-JAN-2000: 2000US-0178572.
PR 02-FEB-2000: 2000US-0179758.
PR 10-FEB-2000: 2000US-0181625.
XX PA
XX (INCY-) INCYTE GENOMICS INC.
PI Baughn MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy R, Lal P,
PI Hillman JL, Azimzai Y, Yue H, Nguyen DB, Yao MG, Gandhi AR,
PI Tang YT, Khan FA:
XX
XX WPI: 2001-418042/44.
DR P-PSDB: AAE04888.
XX
PT Novel human transporter and ion channel proteins useful for treating
PT and preventing transport, neurological, muscle and immunological
PT disorders -
XX
XX Claim 5; Page 140-141; 160pp; English.
PS
XX
XX The present sequence is transporter and ion channel-1 (TRICH-1) cDNA.
CC TRICH is used as vaccine. TRICH is useful for treating a disease or
CC condition associated with decreased expression of functional TRICH,
CC such as transport disorder including amyotrophic lateral sclerosis,
CC cystic fibrosis, Becker's muscular dystrophy, Charcot-Marie-Tooth
CC disease, Duchenne muscular dystrophy, angina and hypertension,
CC neurological disorders including Alzheimer's disease, amnesia, bipolar
CC disorder, dementia, depression, epilepsy, ischemic cerebrovascular
CC disease, stroke, cerebral neoplasms, Pick's disease, Huntington's
CC disease and Parkinson's disease, demyelinating diseases, mental disorder,
CC including mood, anxiety, schizophrenia and seasonal affective disorder,
CC muscle disorder including cardiomyopathy, myocarditis, polymyositis,
CC dermatomyositis, arrhythmias and asthma and immunological disorders
CC including AIDS, adult respiratory distress syndrome (ARDS), allergies,
CC anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
CC syndrome, autoimmune lupus erythematosus and other diseases including
CC sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary
CC artery stenosis, Grave's disease, Cushing's disease, Addison's disease,
CC glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers
CC psoriasis and viral, bacterial, fungal, helminthic and protozoal
CC infections. TRICH DNA is useful in gene therapy and in diagnostic
CC purposes.
XX
XX Sequence 2080 BP; 300 A; 688 C; 632 G; 460 T; 0 other:
XX
XX
XX alignment_scores:
XX
XX Quality: 457.00 Length: 94
XX Ratio: 4.914 Gaps: 0
XX Percent Similarity: 98.936 Percent Identity: 85.106
XX
XX alignment_block:
XX
XX US-09-516-493-12 x AAD09552 ..
XX
XX Align seg 1/1 to: AAD09552 from: 1 to: 2080
XX
XX 1 GluPro1aaSpva1HisLeuGly1LeuAla1TrpLeu1aValGlySer1e 17
XX :::::::::::::::::::: ::::::::::::::::::::
XX 11229 CAGCCTTGTTGATGCCAGCGTGGCGCTGCCCTGGCTGCCTGGGAGCAT 1178
XX
XX 17 tCysLeuPhe1Leu1aGlyPhe1a1aVal1Gly1TrpGly1Pro1Leu1Pro1Trp1 34
XX GGGGCTCTTCATCGCCGGCGTTGGCGTGGCGTGGGCGCCATCCCTGGC 1228
XX
XX 34 euleMetSerGlu1IephProLeuHis1LeuSg1Gly1a1a1aTrnGly 50
XX 12229 TCCTCATGTCAAGAGATCTCCCTTCATATGCAAGGCGCGGACAGGC 1278
XX
XX 51 ValCysVal1LeuTrnAsn1TrpPheMetAla1aPheLeuVal1ThrLysGlu1 67
XX :::::::::::::::::::: ::::::::::::::::::::
XX 1279 ATCTGCGCTCTACCAACTGGCTCATCGCTTCTCTGTGACCAAGAGATT 1328

```

```

67 eassrilemetguilleuAcgpiroTyrGIyAlaphetrlpLeuThra 84
11:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1329 CAGCAGCCTCATGAGGTCTCAgGCCCTATGAGCCTCTGCGTTCCT 1378
84 lAlaphEcysIlleuSerValleuphetrh 94
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1379 CGCCTTTCGCATCTTCAGTCTCCTTTTCACT 1410

seq_name: /SIDS2/gcdata/geneseq/geneseqn/NA2001.DAT:AAF55865

seq_documentation_block:
ID AAF55865 standard; cDNA; 2217 BP.
XX
AC AAF55865:
XX
DT 17-APR-2001 (first entry)
XX
DE Human GLUTX1 coding sequence.
XX
KW Human: GLUTX; gene therapy; vaccine; hexose transport modulator;
KW hexose transport disorder; ischemia; diabetes; hypoglycaemia; ss;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS Homo sapiens.
XX
PN M0200104145-A2.
XX
PD 18-JAN-2001.
XX
PF 14-JUL-2000; 2000MO-IB01042.
XX
PR 14-JUL-1999; 990US-0143907.
PR 27-AUG-1999; 990US-0151140.
PR 23-FEB-2000; 2000US-0184285.
PR 13-JUL-2000; 2000US-0616132.
XX
PA (UYLA-) UNIV LAUSANNE.
XX
PI Thorens B, Ibberson M, Uldry M;
XX
XX WPI: 2001-112615/12.
XX
DR P-PSDB; AAB66932.
XX
PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
PT e.g. ischemia and diabetes -
XX
XX
PS Claim 3; Page 70-71; 124pp; English.
XX
XX The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
XX Cc AAB66932-AA66941). The GLUTX proteins are related to the facultative
XX glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
XX function. The GLUTX proteins may be used in the diagnosis, prevention and
XX treatment of hexose transport disorders such as ischemia, diabetes,
XX hypoglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
XX neurodegenerative disease. The present sequence is the coding sequence
XX for human GLUTX1.
XX
S0 Sequence 2217 BP; 320 A; 750 C; 672 G; 475 T; 0 other;

Alignment_scores:
Quality: 454.00 Length: 94
Ratio: 4.882 Gaps: 0
Percent Similarity: 98.936 Percent Identity: 84.043

alignment_block:
US-09-516-493-1 x AAF55865 ..

Align seg 1/1 to: AAF55865 from: 1 to: 2217

1 GUGPAAAsprValhIsleuGlyleuAlATrpleuAlaValGlyserMe 17
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```
1431 CAGCCTGTTGANGCAGCGGTGGGCTGGCTGGCTGGCCGTGGCAACAT 1480
17 TCysleupheillealaglyphea1aValg1yTrrpGlyProIleProTrpL 34
1481 GTGCCCTTCATGCGCGGCTTTTCCGGTGGGCTGGGGCCCATGCCCTGGC 1530
34 euleuMetsergluilepheProleuHis1leuSglYala1aThrngly 50
1531 TCCTCATGTACAGATCTTCCTCTGCATGTCAAGGCGCGTGGACAGGC 1580
51 ValcysValleuthrasnTrpMetAlaPheLeuValThrlySgluph 67
1581 ACTCGGCTTCACCAACTGGCCATGAGCGCTTTCGTGACCAAGAGATT 1630
67 eaasSerIleMetgluileuArpProTrrgYalAlaPheTrpLeuThra 84
1631 CAGACGCTCATGAGAGCTCTCAGGCCCTTATGAGCCCTTCTGGCTTGCT 1680
84 lAlaPheCysIleLeuSerValleuPheThr 94
1681 CGCCTTTCGATCTTCAGTGTCTTTTCACT 1712

seq_name: /SID52/gcdata/geneseq/geneseqn/NA2000.DAT:AAA44914

seq_documentation_block:
ID AAA44914 standard: cDNA: 579 BP.
XX
AC AAA44914:
XX
DT 21-AUG-2000 (first entry)
XX
DE Human secreted expressed sequence tag SEQ ID NO:1489.
XX
KW Human: mouse; chicken; rat; secreted expressed sequence tag; SEST;
expressed sequence tag; EST; probe; chemotactic; proliferative;
immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
antiviral; antidiabetic; antisthmatic; vulnery; antiparkinsonian;
antitumor; osteopathic; neuroprotective; neurotropic; antipsoriatic;
cerebroprotective; anticonvulsant; antidepressant; gene therapy;
vaccine; autoimmune diabetes; multiple sclerosis; allergic condition;
insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
central nervous system disorder; Alzheimer's disease; stroke;
Parkinson's disease; Huntington's disease; coagulation disorder;
haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
tumour; infection; depression; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN W0200021991-A1.
XX
PD 20-APR-2000.
XX
PF 15-OCT-1999; 99WO-US24206.
XX
PR 15-OCT-1998; 98US-0104436.
XX
PA (GEMX ) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Bowman MR;
XX
DR WPI: 2000-317938/27.
XX
PT Isolated polynucleotides, and encoded proteins, comprising secreted
PT expressed sequence tags (SESTs), useful for treating various disorders
PT such as autoimmune, infectious, and central nervous system disorders -
XX
PS Claim 1; Page 586; 803pp; English.
XX
CC AAA43426 to AAA45925 represent specifically claimed secreted expressed
CC sequence tags (SESTs), isolated from human, mouse, chicken and rat
```

```
CC tissue sources. The SESTs can have a range of activities depending on
CC the tissues they were isolated from. The activities include:
CC chemotactic; proliferative; immunomodulatory; haematopoietic;
CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
CC antisthmatic; antitumor; antitumor; antisthmatic; neuroprotective;
CC neurotropic; antiparkinsonian; antipsoriatic; cerebroprotective;
CC anticonvulsant; and antidepressant. The SESTs can be used for gene
CC therapy and in vaccines. The SESTs are useful as probes for the
CC identification and isolation of full-length cDNAs and genomic DNA
CC molecules which correspond to the SESTs. Proteins encoded by the SESTs
CC are useful in assays for determining biological activity and raising
CC antibodies. They may be useful for treatment of autoimmune disorders
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC osteoporosis, osteoarthritis, central nervous system disorders
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC disease), tumours, bacterial, fungal or viral infections, depression and
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
CC in the exemplification of the present invention.
XX
SQ Sequence 579 BP; 109 A; 184 C; 180 G; 106 T; 0 other;
```

```
alignment_scores:
Quality: 266.50 Length: 70
Ratio: 4.675 Gaps: 1
Percent Similarity: 81.429 Percent Identity: 72.857
```

alignment_block:

US-09-516-493-12 x AAA44914/rev ..

Align seg 1/1 to reverse of: AAA44914 from: 1 to: 579

```
2 ProAlaSPValHisLeuGlyLeuAlaTrpLeuAlaValGlySerMetCys 18
18 TCysleupheillealaglyphea1aValg1yTrrpGlyProIleProTrpLeu 35
364 TCCTCATGTACAGATCTTCCTCTGCATGTCAAGGCGCGTGGACAGGCATC 265
393 CCGCGGACCGCTGATCGCTGTCTTGGC.....TC 365
18 sleupheillealaglyphea1aValg1yTrrpGlyProIleProTrpLeu 35
364 TCCTCATGTACAGATCTTCCTCTGCATGTCAAGGCGCGTGGACAGGCATC 265
35 euleuMetsergluilepheProleuHis1leuSglYala1aThrnglyVal 51
314 TCATGTACAGATCTTCCTCTGCATGTCAAGGCGCGTGGACAGGCATC 265
52 CysValleuthrasnTrpMetAlaPheLeuValThrlySglupheas 68
264 TGGGTCTTCACCAACTGGCCATGAGCGCTTTCGTGACCAAGAGATTGAG 215
68 nserIleMet 71
214 CAGCCTTCATG 205
```

seq_name: /SID52/gcdata/geneseq/geneseqn/NA2001.DAT:AAF67134

seq_documentation_block:

ID AAF67134 standard: cDNA: 378 BP.

```
XX
AC AAF67134:
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 2890.
XX
KW Human: cytosolic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN W0200102568-A2.
```

PD 11-JAN-2001.
 XX
 PF 30-JUN-2000: 2000WO-US18374.
 XX
 PR 02-JUL-1999: 99US-0142310.
 PR 02-JUL-1999: 99US-0142311.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEO INC.
 XX
 PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
 PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
 PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
 PI Kita D, Garcia V, Jones LW, Strache-Grain B;
 XX
 WPI: 2001-091805/10.
 XX
 PT library of polynucleotides for diagnosing a cancerous state of a
 PT mammalian cell and detecting cancer, particularly of the colon or
 PT prostate, comprises 3351 human polynucleotide sequences -
 XX
 PS Claim 9: Page 974: 1046pp: English.

XX The present sequence is one of 3351 sequences in a library of human
 CC polynucleotides. The library is used to detect differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell and can
 CC detect colon, prostate, breast and lung cancer. The library can be used
 CC to produce probes for detection of mRNA and to produce additional copies
 CC of the polynucleotides. The probes can be used for chromosome mapping of
 CC the polynucleotide and for detection of transcription levels. Ribozymes
 CC or antisense oligonucleotides can be generated. The polynucleotides and
 CC their gene products are used as genetic or biochemical markers (e.g. in
 CC blood or tissues) that will detect the earliest changes along the
 CC carcinogenesis pathway and/or monitor the efficacy of therapies and
 CC preventive interventions. The polynucleotides, polypeptides and
 CC antibodies against them can be used in pharmaceutical compositions to
 CC treat the cancers and proliferative disorders such as neoplasia,
 CC dysplasia and hyperplasia.
 XX
 SO Sequence 378 BP; 45 A; 130 C; 117 G; 86 T; 0 other:

alignment_scores: Quality: 221.00 Length: 77
 Ratio: 3.348 Gaps: 1
 Percent Similarity: 85.714 Percent Identity: 59.740

alignment_block:
 US-09-516-493-12 x AAF67134 ..

Align seg 1/1 to: AAF67134 from: 1 to: 378

19 LeuPhel1eAlaGlyPheAlaValGlyTrpPolypPro1lePrrPleuLe 35
 ||||||| ||||| ||||||| ||||||| ||||||| ||||||| |||||||
 31 CTCCTTCAACATGAGGCTACGCGCTGGGCTGCCATCCACTGCGCTCT 80
 35 uMeSecl1u1lePheProleuHis1leLysGlyValAlaIarhrgYVa1C 52
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 81 CATGTGTAGTCTCGCCCTGCGCGTGGCGCTGGCGCTCAAGGACTCT 130
 52 ySVal1eUThrAsnTrPPhMeLa1aPheLeValThrLysGluPheAsn 68
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 131 GCGTGTGTCGACGCTACGCGCTTCGTCACCAAGCTCTTCCTG 180
 69 Ser1eMeGlu1leUarPrrTrGlyAlaPheTrPleuThrAlaA 85
 ||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 181 CCACTGTG. GAGCACTTCGCGCTCCAGTGCCTTCTTCTTCTTCCCGG 229
 85 lApheCys1leUeSerVal1eUpherThr 94
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 230 CCATCTGCTTGTGAGCTGCTGCTTTCACA 258

seq_name: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT:AAA23461

seq_documentation_block:
 ID AAA23461 standard: cDNA: 823 BP.
 XX
 AC AAA23461:
 XX
 DT 19-JUN-2000 (first entry)
 XX
 DE cDNA encoding human secreted protein vq1_1, SEQ ID NO:77.
 XX
 KW Human: secreted protein; cancer; tumour; cardiovascular disorder;
 KW blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;
 KW infection; fungal; bacterial; viral; HIV; allergy; arthritis;
 KW neurodegenerative disease; asthma; contraceptive; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 54..371
 FT /*tag= a
 FT /product= "Human secreted protein vq1_1"
 XX
 PN W0200011015-A1.
 XX
 PD 02-MAR-2000.
 XX
 PF 24-AUG-1999: 99WO-US19351.
 XX
 PR 24-AUG-1998: 98US-0097638.
 PR 24-AUG-1998: 98US-0097659.
 PR 09-SEP-1998: 98US-0099618.
 PR 28-SEP-1998: 98US-0102092.
 PR 25-NOV-1998: 98US-0109978.
 PR 23-DEC-1998: 98US-0113645.
 PR 23-DEC-1998: 98US-0113646.
 PR 23-AUG-1999: 99US-0379246.
 XX
 PA (ALPH-) ALPHAGENE INC.
 XX
 PI Valenzuela D, Yuan O, Hoffman H, Hall J, Raplejo P;
 DR WPI: 2000-224657/19.
 DR P-PDB: AAY95019.
 XX
 PT New secreted or transmembrane proteins and polynucleotides encoding
 PT them, useful for treating neurodegenerative disorders, autoimmune
 PT diseases and cancer -
 XX
 PS Claim 86: Page 335: 357pp: English.
 XX
 CC The invention relates to 40 human secreted proteins (AAY94981-Y95020),
 CC and cDNA sequences encoding them (AAA23423-A22462). The secreted
 CC proteins of the invention include those that are thought to be only
 CC partially secreted, i.e., transmembrane proteins. The proteins of the
 CC invention may exhibit one or more activities selected from the following:
 CC cytokine activity; cell proliferation; differentiation; immune
 CC modulation; haematopoiesis regulation; tissue growth activity;
 CC activin/inhibin activity; chemotactic/chemokinetic activity; haemostatic
 CC and thrombolytic activity; anti-inflammatory activity; and tumour
 CC inhibition activity. The proteins may be administered to patients as
 CC vaccines, and the nucleotides may be used as part of a gene therapy
 CC regime. Diseases or conditions that may be treated using the proteins or
 CC nucleotides of the invention include autoimmune diseases; genetic
 CC disorders; haemophilia; cardiovascular diseases; cancer; bacterial,
 CC fungal and viral infections, especially HIV; multiple sclerosis;
 CC rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome;
 CC insulin dependent diabetes mellitus; and allergic reactions such as
 CC asthma and anaemia. They may also be used for treating wounds, burns,
 CC ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease and amyotrophic
 CC lateral sclerosis (ALS). Proteins with activin/inhibin activity may
 CC additionally be useful as contraceptives. Nucleic acid sequences of the
 CC invention may be used in chromosome mapping, and as a source of

CC diagnostic primers and probes. The present sequence represents cDNA
 CC encoding one of the 40 proteins of the invention.
 XX
 SO Sequence 823 BP; 170 A; 257 C; 233 G; 163 T; 0 other;

alignment_scores:
 Quality: 221.00 Length: 77
 Ratio: 3.348 Gaps: 1
 Percent Similarity: 85.714 Percent Identity: 59.740

alignment_block:
 US-09-516-493-12 x AAA23461 ..

Align seg 1/1 to: AAA23461 from: 1 to: 823

```

19 LeupheIIAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeu 35
   |||||  |||:|||||:|||||:|||||  |||||
57 CTTCTCATGATGGGCTACGCCGGGCTGGGCTCCATCAGCTGGCTGCT 106
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
35 uMetSerGluIlePheProLeuHisIleLysGlyValAlaIleThrGlyValC 52
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
107 CATGCTGTAGAGTCTCCCTCCCTGGCTGCCCTGGCTCCAGGCTCT 156
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
52 yValLeuThrAsnTrpPheMetAlaPheLeuValThrLysGluPheAsn 68
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
157 GCGTGTGCGCCAGCTGCTCACCGCTTCTCTCCACCAAGTCTTCTG 206
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
69 SerIleMetGluIleLeuArgProTyrGlyAlaPheTrpLeuThrAlaA 85
   |||||  |||  |||||  |||  |||||  |||  |||||
207 CCAAGTGTGTGAGCACCTTGGCTCCAGGCTGCTTCTTCTTCTGCGG 255
   |||||  |||||  |||||  |||||  |||||  |||||
85 IapheCysIleLeuSerValLeuPheThr 94
   ||:||||:||||:||||:||||:||||:
256 CCATCTGCTTGGTGAAGCCGTGTTCACCA 284

```

seq_name: /SIDS2/gcdata/geneseq/geneseq/NA2000.DAT: AAC80569

seq_documentation_block:

ID AAC80569 standard; cDNA; 1577 BP.

XX AAC80569;
 AC
 DT 12-FEB-2001 (first entry)
 XX
 DE Human secreted protein gene 39 SEQ ID NO:49.
 XX
 KW Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;
 KW antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective;
 KW neotropic; neuroprotective; antibacterial; virucide; fungicide; cancer;
 KW ophthalmological; autoimmune disease; hyperproliferative disorder;
 KW cardiovascular disorder; cerebrovascular disorder; wound healing;
 KW nervous system disorder; aging; chemotaxis; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200058467-A1.
 XX
 PD 05-OCT-2000.
 XX
 PF 22-MAR-2000; 2000WO-US07505.
 XX
 PR 26-MAR-1999; 99US-0126502.
 PR 17-DEC-1999; 99US-0172410.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Komatsoulis G;
 XX
 DR WPI: 2000-611712/58.
 DR P-PSDB: AAB45158.
 XX
 PT Nucleic acid molecules encoding human secreted proteins, used in

PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 PT Parkinson's diseases and cancers -
 XX
 PS Claim 1: Page 371-372; 440pp; English.

XX Polynucleotide sequences AAC80531-C80580 represent cDNA encoding human
 CC secreted proteins AAB45120-B45169. Sequences AAB45170-B45225 represent
 CC alternative polypeptides encoded by the genes, and amino acid sequences
 CC to which they are homologous. The genes and proteins have activities
 CC dependent on the tissues and cells in which they are expressed. Examples
 CC of their activities include immunosuppressive; antirheumatic;
 CC antineoplastic; antiproliferative; cytostatic; cardiac; vasotropic;
 CC cerebroprotective; neuroprotective; neuroprotective; antibacterial; virucide;
 CC fungicide; and ophthalmological. The secreted proteins, polynucleotides,
 CC antagonists and agonists may be useful in treating, preventing and/or
 CC diagnosing diseases and disorders such as autoimmune diseases
 CC e.g., rheumatoid arthritis, hyperproliferative disorders e.g., neoplasms
 CC of the breast or liver, cardiovascular disorders e.g., cardiac arrest,
 CC cerebrovascular disorders e.g., cerebral ischaemia, angiodysgenesis, nervous
 CC system disorders e.g., Alzheimer's disease, infections caused by
 CC bacteria, viruses and fungi and ocular disorders e.g., corneal infection.
 CC The polypeptides can also be used to aid wound healing and epithelial
 CC cell proliferation, to prevent skin aging due to sunburn, to maintain
 CC organs before transplantation, for supporting cell culture of primary
 CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
 CC also be used as a food additive or preservative to increase or decrease
 CC storage capabilities. AAC80522-C80530 and AAB45119 represent sequences
 CC used in the isolation and characterisation of the genes and proteins of
 CC the invention.
 CC
 SO Sequence 1577 BP; 352 A; 496 C; 445 G; 284 T; 0 other;

alignment_scores:
 Quality: 221.00 Length: 77
 Ratio: 3.348 Gaps: 1
 Percent Similarity: 85.714 Percent Identity: 59.740

alignment_block:
 US-09-516-493-12 x AAC80569 ..

Align seg 1/1 to: AAC80569 from: 1 to: 1577

```

19 LeupheIIAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeu 35
   |||||  |||:|||||:|||||:|||||  |||||
291 CTTCTCATGATGGGCTACGCCGGGCTGGGCTCCATCAGCTGGTGT 340
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
35 uMetSerGluIlePheProLeuHisIleLysGlyValAlaIleThrGlyValC 52
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
341 CATGCTGTAGAGTCTCCCTCCCTGGCTGCCCTGGCTCCAGGCTCT 390
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
52 yValLeuThrAsnTrpPheMetAlaPheLeuValThrLysGluPheAsn 68
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
391 GCGTGTGCGCCAGCTGCTCACCGCTTCTCTCCACCAAGTCTTCTG 440
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
69 SerIleMetGluIleLeuArgProTyrGlyAlaPheTrpLeuThrAlaA 85
   |||||  |||  |||||  |||  |||||  |||  |||||
441 CCAAGTGTGTGAGCACCTTGGCTCCAGGCTGCTTCTTCTTCTGCGG 489
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
85 IapheCysIleLeuSerValLeuPheThr 94
   ||:||||:||||:||||:||||:||||:
490 CCATCTGCTTGGTGAAGCCGTGTTCACCA 518

```

seq_name: /SIDS2/gcdata/geneseq/geneseq/NA2001.DAT: AAF55870

seq_documentation_block:

ID AAF55870 standard; cDNA; 1541 BP.

XX AAF55870;
 AC
 DT 17-APR-2001 (first entry)
 XX
 DE Human GLUTX3 coding sequence.


```

XX Human; GLUTX: gene therapy; vaccine; hexose transport modulator;
KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia; SS;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX Homo sapiens.
XX MO200104145-A2.
XX PN 18-JAN-2001.
XX PD 14-JUL-2000; 2000MO-1B01042.
XX PF 14-JUL-1999; 99US-0143907.
XX PR 27-AUG-1999; 99US-0151140.
XX PR 23-FEB-2000; 2000US-0184285.
XX PR 13-JUL-2000; 2000US-0616132.
XX PA (UYLA-) UNIV LAUSANNE.
XX P1 Thorens B, Ibberson M, Uldry M;
XX DR WPI: 2001-112615/12.
XX DR P-PSDB; AAB66937.
XX PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
XX PM the prevention, diagnosis and treatment of hexose transport disorders,
XX PS e.g. ischemia and diabetes -
XX CC Claim 3; Page 81-82; 124pp; English.
XX SC Sequence 1541 BP; 209 A; 536 C; 482 G; 311 T; 3 other;
XX SQ
Alignment_scores:
Quality: 219.00 Length: 77
Ratio: 3.318 Gaps: 1
Percent Similarity: 85.714 Percent Identity: 59.740
alignment_block:
US-09-516-493-12 x AAF55870 ..
Align seg 1/1 to: AAF55870 from: 1 to: 1541
19 leupheIIlealagIphealaIvalIGlYTRgIyProtllePOTrPlleue 35
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1220 CTtTtATAtCATGCGGTACGCCGTCGGTGCGGTCCATCAGCTGGTCCT 1269
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
35 uMetSerGIuIllePhleProleuhIslelySGlyValAlaThrlglYAlc 52
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1270 CAGTGTGCAGAGCTCAACCCTGCCTGCCGTGCCGTGCGCCACAGGCGCT 1319
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
52 ySValleuTrAsnTrPhMeTaIaphleuVaIlThrlySGluPhasn 68
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1320 GCCTGCTGGACTGGCTCACCGCCTTGCTGCCTCACCACTCCCTCTG 1369
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
69 SerIlleMeTGlulleuArGPOTrYrGLyAlaPhetRpleuthr.AlaA 85
::: ||| ||||| ||||| ||||| ||||| ||||| |||||
1370 CcAGTGGT.GAGCACCTTCGGCCTCcAGGrGCCCTTCCTTCCTTCGGGG 1418
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
85 laphecysIlleleuSerVallleuPherthr 94
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1419 cCATTCcCTcCTGcGTGAGCcCTGcGTTCACAA 1447

```

```

seq_name: /SIBS2/gcgcdata/geneseq/geneseqn/NA2001.DAT:AAF55871
seq_documentation_block:
ID   AAF55871 standard; cDNA; 2011 BP.
AC   AAF55871;
DT   17-APR-2001 (first entry)
DE   Rat GLUTX3 coding sequence.
XX
KW   Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;
KM   hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
KN   hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
OS   Rattus sp.
MO   Z000104145-A2.
XX
PD   18-JAN-2001.
XX
PF   14-JUL-2000; 2000WO-IB01042.
XX
PR   14-JUL-1999; 99US-0143907.
PR   27-AUG-1999; 99US-0151140.
PR   23-FEB-2000; 2000US-0184285.
PR   13-JUL-2000; 2000US-0616132.
XX
PA   (UTLA-) UNITV LAUSANNE.
XX
PI   Thorens B, Ibberson M, Uldry M;
XX
DR   WPI: 2001-112615/12.
DR   P-PSDB; AAB66938.
XX
PT   Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PP   the prevention, diagnosis and treatment of hexose transport disorders,
PT   e.g. ischemia and diabetes -
XX
PS   Claim 3; Page 82-83; 124pp; English.
XX
CC   The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC   AAB66932-AAB6941). The GLUTX proteins are related to the facilitative
CC   glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC   function. The GLUTX proteins may be used in the diagnosis, prevention and
CC   treatment of hexose transport disorders such as ischaemia, diabetes,
CC   hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC   neurodegenerative disease. The present sequence is the coding sequence
CC   for rat GLUTX3.
XX
SO   Sequence 2011 BP; 337 A; 621 C; 593 G; 459 T; 1 other;

alignment_scores:
      Quality: 219.00          Length: 94
      Ratio: 3.042            Gaps: 1
Percent Similarity: 76.596    Percent Identity: 52.128

alignment_block:
US-09-516-493-12 x AAF55871 ..
Align seg 1/1 to: AAF55871 from: 1 to: 2011

2 ProAlaSPValHisIleuGlyLeuAlaTrpLeuAlaValGlySerMetCys 18
||||| ::: ||||::: ::::: :::::
1153 CCAGCGCGGGCTTCAACTATCTTCACCTGTACCCTCGTGCCACCATT 1202

18 sleuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProtrrPeul 35
||||||| |||:::|||||:::||||| |||||
1203 GCCTTCATTATGGCGCATGCATCGGCTGGGGGCCCATCACCTGGCTCC 1252

35 emueTserrGuilePheProLeuHisIleLysGlyValAlaThrGlyVal 51
|||||:::|||||::: |||||:::|||||:::

```

```
1253 TCATCTCTGAGGTCTGCCCCGTCGCGCCGTCGTGCGCTCAGGCGTC 1302
      52 CysValleuThrAsnTriphemeAlaPheleuValThrLysGluPheas 68
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1303 TCGCTGCTGGTAGCTGCTGCTCAGCCCTTCGTCCTCACAAGTACTTCT 1351
      68 nSerIleMetGluIleLeuArgProTyrGlyAlaPheTrpLeuThr.Ala 84
      ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1352 .GCTGCAGTGAATGCTTCGCGCTCCAGTGCCTTCTTCTTCTTCTCTCG 1401
      85 AlaPheCysIleLeuSerValleuPheThr 94
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1402 GCCATCTGCTGCTCAGCCCTCCTTTCACA 1431
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAC40459
seq_documentation_block:
ID AAC40459 standard; DNA: 1755 BP.
XX
AC AAC40459;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 28375.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
```

```
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
```



```
XX Tang YT, Liu C, Drmanac RT;
PI WPI: 2001-457603/49.
DR P-PSDB: AAM25722.
XX
PT Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
XX Claim 1: Page 564: 1217pp: English.
XX
CC AAM25963 to AAM25964 encode the human proteins given in AAM25925 to
CC AAM25963. The proteins can have activities based on the tissues and
CC cells they are expressed in, such as: antiinflammatory; antineuritic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antilaemic; antiaggregant; haemostatic; vulnerary;
CC antilicer; osteopathic; dermatoprotective; antiallergic; antilastmatic;
CC antidiabetic; cytosstatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopenia, wounds, allergic
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders.
XX
SQ Sequence 421 BP: 72 A: 140 C: 122 G: 87 T: 0 other:

alignment_scores:
    Quality: 189.00      Length: 56
    Ratio: 3.780        Gaps: 0
Percent Similarity: 89.286   Percent Identity: 57.143

alignment_block:
US-09-516-493-12 x AAM99663 ..

Align seg 1/1 to: AAM99663 from: 1 to: 421

12 LeuAlaValAlGlySerMetCysLeuPheIleAlaGlyPheAlaValAlGlyTr 28
|||||
1 CTGCGAGCGCCGCAAGTCTGCTCTTCGTCAGCGCTACGTCGCGCGTGG 50
28 PGLyProIleProTrPleuLeuMetSerGluIlePheProIleuHsIleL 45
|||||
51 GGGTCCCATCACCTGCTGCTGATGTCAGGCTCTGCCCGCCGCGGCC 100
45 ysgIyValAlaIleThiGlyValCysValIleuThrAsnTrPheMetAlaPhe 61
:::|||||
101 GTGGCGGTGGCCCTCAGGCGCTGCTGCGGCGCAGCTGCTACCGCCTTC 150
62 LeuValThrLysGluPhe 67
:::|||||
151 GTCTCCACCAAGTCTCTTC 168

seq_name: /SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAC45857
seq_documentation_block:
ID AAC45857 standard; DNA: 1395 BP.
XX
XX AAC45857;
AC
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 48015.
XX
```

```
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
```


72 uilleuAArgpToTYrGlyAlaPheTrpIeuThraAlaAlaPheCysIleL 89
: |||||::||| :|||::|
1266 C...TGAGACTCTTACGAGACTTCTCATTACGCTGCATCAGCGC 1312
89 euserValleuphe 93
||:::|||||
1313 TGGCCATTGCTTTC 1326

seq_name: /SID52/gcgdata/geneseq/geneseqn/NA2000.DAT: AAC39099
seq_documentation_block:
ID AAC39099 standard; DNA: 1683 BP.
XX
AC AAC39099;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23364.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 28-APR-1999; 990S-0130891.
PR 30-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 04-MAY-1999; 990S-0132407.
PR 05-MAY-1999; 990S-0132484.
PR 06-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.

PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 02-JUL-1999; 990S-0142154.
PR 06-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 28-JUL-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.

```
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

```
alignment_scores:
    Quality: 187.50
    Ratio: 2.799
    Percent Similarity: 76.136
    Length: 88
    Gaps: 1
    Percent Identity: 43.182
```

```
alignment_block:
    US-09-516-493-12 x AAC39099
```

```
Align seg 1/1 to: AAC39099 from: 1 to: 1683
6 HistLeuGlyLeuAlaIrrpleuAlaValGlyserMeCysLeuPheIleAl 22
   ||| : : : : : ||||| : : : : : : : : : : : : : : :
1185 CACGAGACAGCTCCAGCTCCTGCTGTTGTTGTAATATGATGTTACAG 1234
22 acylPheAlaValAlGlyTrpGlyProIleProIrrpleuLeuMeSerGluI 39
   : : : : : : : : : : ||| : : : : : : : : : : : : : :
1235 ATCGTTTTCAGCAGGAATGAGACATGCCCTGCGTGGTCATCTCGA 1284
39 IePheProLeuHsIleLysGlyValAlaThrGlyValCysValLeuThr 55
   ||||| : : : : : ||||| : : : : : : : : : : : : : :
1285 TATTCCCATTAACATPAAAGAGTAGCAGAGCATGCGACGCTGTCG 1334
56 AsnTrpPheMeAlaPheLeuValThrLysGluPheAsnSerIleMetG 72
   ||||| : : : : : : : : : : ||||| : : : : : : : : :
1335 AACTGTTTGGAGCGTGGGCTGTTCTTACACTTTCACCTTCCTCAATG 1384
72 uIleLeuArgProTyrglyAlaPheTrpPleuThrAlaAlaPheCysIleL 89
   : : : : : : : : : : ||||| : : : : : : : : : : : :
1385 C...TGGAGCTCTTACGGAACCTTCTCATTTACGCTGCCATCAAGCCG 1431
1432 TGSCCATTTGCTTC 1445
89 euSerValLeuPhe 93
   || : : : : : : : : :
1385 C...TGGAGCTCTTACGGAACCTTCTCATTTACGCTGCCATCAAGCCG 1431
1432 TGSCCATTTGCTTC 1445
seq_name: /SID2/gcdata/geneseq/geneseq/NA2000.DAT:AAC45298
seq_documentation_block:
ID AAC45298 standard; DNA: 1685 BP.
XX
AC AAC45298;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 46007.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132487.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
```

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144864.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149724.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0152707.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159296.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.


```

55  ThisAntpPheMetAlaPheLeuValThrIysGluPheAsnSerIleW 71
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
1856  GCGAAGCTGATGGCCGATTCATCGTGGACACATTTTCCGATGTTAA 1905
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
71  tGlutIleuArgProTyrGly...AlaPheTrpLeuThrAlaIlePhe. 86
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
1906  CGATTGTGGGGCGGACGAGCAAGATCTACTGATCTTTCGACGCGATGG 1955
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
87  ..CysIleLeu 89
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
1956  CCGTGTCTCTC 1966

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-591-025-8

seq_documentation_block:
; Sequence 8, Application US/09591025
; Patent No. 6303373
; GENERAL INFORMATION:
; APPLICANT: Bogan, Jonathan S.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: Method of Measuring Plasma Membrane
; TITLE OF INVENTION: Targeting of GLUT4
; FILE REFERENCE: 0399.1210-004
; CURRENT APPLICATION NUMBER: US/09/591.025
; CURRENT FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/154,078
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/138,237
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 2592
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified GLUT4 containing myc tag sequences
US-09-591-025-8

alignment_scores:
    Quality: 120.50      Length: 86
    Ratio: 2.152        Gaps: 1
    Percent Similarity: 65.116      Percent Identity: 29.070

alignment_block:
US-09-516-493-12 x US-09-591-025-8 ..
Align seg 1/1 to: US-09-591-025-8 from: 1 to: 2592

9  LeuAlaTrpLeuAlaValAlGlySerMetCysLeuPheIleAlaGlyPheAl 25
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
1453  ATGAGCTACGTCATTCATTCGTCATCTTGGCTTCGTCATCTTGTGA 1502
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
25  aValGlyTrpGlyProIleTrpLeuLeuMetSerGluIlePheProL 42
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
1503  GATTGGCCGCGCCCATTCCTTGGTTCATCGTGGCCGAGGCTTTAGCC 1552
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
42  euNIleIleLysGlyValAlaThrGlyValCysValLeuThrAsnTrpPhe 58
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
1553  AGGAGACCCCGCCGAGCCATGCGCTGCTGTTTCTCCAACTGGAGG 1602
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
59  MetAlaPheLeuValThrIysGluPheAsnSerIleMetGluIleLeuAr 75
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
1603  AGCAACTTCATTCATTCGCGATGGGTTTCCAGTATGTTGCGGAGGCTATGGG 1652
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
75  gProTyrGlyAlaPheTrpLeuThrAlaIlePheCysIleLeuSerValL 92
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
1653  GGCCTAAC...GTCTTCCTTCTATTTGGGCTCTCTGCTGGGCTTGTCA 1699
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
92  euPheThr 94
    :::::::::::::::::::::|||||:|||||:|||||:|||||:

```

```

1700  TCTTCACC 1707

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-08-998-416-719

seq_documentation_block:
; Sequence 719, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998.416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CCCI976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 719:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 708 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAGI471RP
US-08-998-416-719

alignment_scores:
    Quality: 100.00      Length: 77
    Ratio: 1.887        Gaps: 2
    Percent Similarity: 68.831      Percent Identity: 29.870

alignment_block:
US-09-516-493-12 x US-08-998-416-719/rev ..
Align seg 1/1 to reverse of: US-08-998-416-719 from: 1 to: 708

18  CysLeuPheIle..AlaGlyPheAlaValAlGlyTrpGlyProIleProTrpL 34
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
536  TGCCTCTACTATTTTCTCTTTCACACTACCTGCGCCCATTCGCTATG 487
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
34  euLeuMetSerGluIlePheProLeuNIleIleLysGlyValAlaThrGly 50
    :::::::::::::::::::::|||||:|||||:|||||:|||||:
486  TCATTGTTCGAGACTTTCATTAAGACTCAAGCCAGGCGCATGGCT 437
    :::::::::::::::::::::|||||:|||||:|||||:|||||:

```

[illegible]

```

1 seq.documental.alignment_block:
2 Sequence 1, Application US/09031392
3 Patent No. 5942398
4 GENERAL INFORMATION:
5 APPLICANT: Tartaglia, Louis A.
6 APPLICANT: Meng, Xun
7 TITLE OF INVENTION: NUCLEIC ACID MOLECULES
8 TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
9 NUMBER OF SEQUENCES: 10
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: Fish & Richardson P.C.
12 STREET: 225 Franklin Street
13 CITY: Boston
14 STATE: MA
15 COUNTRY: USA
16 ZIP: 02110-2804
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Diskette
19 COMPUTER: IBM Compatible
20 OPERATING SYSTEM: Windows95
21 SOFTWARE: FastSeq for Windows Version 2.0
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/09/031,392
24 FILING DATE: 26-FEB-1998
25 ATTORNEY/AGENT INFORMATION:
26 NAME: Melkielejohn, Ph.D., Anita L.
27 REGISTRATION NUMBER: 35,283
28 REFERENCE/DOCKET NUMBER: 07334/072001
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: 617/542-5070
31 TELEFAX: 617/542-8906
32 TELEX: 200154
33 INFORMATION FOR SEQ ID NO: 1:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 2343 base pairs
36 TYPE: nucleic acid
37 STRANDEDNESS: single
38 TOPOLOGY: linear
39 MOLECULE TYPE: cDNA
40 FEATURE:
41 NAME/KEY: Coding Sequence
42 LOCATION: 73...1761
43 US-09-031-392-1
44
45 alignment_scores:
46 Quality: 84.50 Length: 83
47 Ratio: 1.625 Gaps: 1
48 Percent Similarity: 62.651 Percent Identity: 28.916
49
50 alignment_block:
51 US-09-516-493-12 x US-09-031-392-1 ..
52
53 Align seg 1/1 to: US-09-031-392-1 from: 1 to: 2343
54
55 11 TTPLEUAlaVala(G)SermetycysLeuPheIIeAlaGlyPheAlaIgl 27
56 ::::::::::::::::::::::::::::::::::::::::::::::::::::
57 1309 TACCTGAGATGATGGTGCGCAATTCGGCCATCATCGCCTCTTCTGCACTGG 1358
58
59 27 YTPRLcylProIIleProTRPLeuLeuMetSerGluIIlePheProLeuHis 44
60 I IIII ::::::::::::::::::::
61 1359 GCGAGGTGGCAATCCGCTTCATCTTGACGTGTGAGTTCTTCCAGCAATCTC 1408
62
63 44 IeIySGIYAlaIaThrGlyValaCysValleuThrAsnTRPPIleMetAla 60
64 ::::::::::::::::::::
65 1409 AGCGGCGCGCTCCCTTCATCATTTGACAGGCAACCTCAACTGCTCTCCAAC 1458
66
67 61 PheIeuValaThrLysIleuPheAsnSerIIleMetGluIIleLeuArgProTy 77
68 IIII ::::::::::::::::::::
69 1459 TTTGGCTGTGGGCTCTCTTCCCATTCATTCAGAAAGTGTGGACACACTA 1508
70
71 TGIYAlaPheTRPLeuThrAlaIaIaPheCysIIleLeuSerValleuPhe 93

```



```
191 ACATACGACAAACAGCTGGGATATGCGACAGAAACATGACCAATT 240
32 .....Profrpleu 34
241 TCACATAAAGCTTGGTCATACAAAGTTCTTGTTAGTGAGCTGGAG 290
35 LeuMetSerGluIlePheProLeuHisIleYsgIyValAlaThrClYVa 51
291 ATGCCAAGAGCTGTTTCACCTGCATGACAAAGCATTTTTCACCAACC 340
51 lCysValLeuThrAsnTrpPheMetAlaPheLeuValThrIysGluPheA 68
341 CTGTGTGCAGCCTCAAGCTAATGGC.....TACA 372
68 snSerlIleMetGluIleLeuArgProYrGlyAlaPheTrp 81
373 ACTATGCCATGTTTGGCTTCACCTCTTATGTCCTTATTGG 413
seq_name: /cgn2_6/ptodata/2/1na/backfiles1.seq:5198359-3
seq_documentation_block:
; Patent No. 5198359
; APPLICANT: TANIGUCHI, TADATSUGU; JATAKEYAMA, MASANORI;
; MINAMOTO, SEIJIRO; KONO, TAKESHI; DOI, TAKESHI; MIYASAKA, MASAYUKI;
; TSUDO, MITSURU; KARASUYAMA, HAJIME
; TITLE OF INVENTION: RECOMBINANT PROTEIN RECEPTOR FOR IL-2
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/487,059
; FILING DATE: 05-MAR-1990
; SEQ ID NO:3:
; LENGTH: 2306
5198359-3
```

```
alignment_scores:
  Quality: 74.50      Length: 71
  Ratio: 1.961      Gaps: 5
  Percent Similarity: 53.521      Percent Identity: 29.577
```

```
alignment_block:
US-09-516-493-12 x 5198359-3 ..
```

```
Align seg 1/1 to: 5198359-3 from: 1 to: 2306
```

```
5 ValHisLeu...GlyLeuAlaTrpLeuAla.....GlyP 13
|||||
2074 GTCCACTGCTGAGGAGCTGCTCGCTATCCCTGCCCCCCTCCAAACT 2123
14 .....ValGlySerMetCysLeuPheIleAla.....GlyP 24
2124 TAATCATCCACTTCTGAACCTCATTTGCTACTTCCGTGGCTAACAGGGT 2173
24 heAlaValAlGlyTrpGly.....ProIleProTrp 33
|||||
2174 TTGGTGGAGGGTGGGAGGAGCGGTGCTCAGCTCCACTGCCCTAT 2223
34 LeuLeuMetSerGluIlePheProLeuHisIleYsgIyValAlaThrCl 50
|||||
2224 TTAGTCATG.....AGCTCAGTACGTTCCG 2249
50 yValCysValLeu 54
2250 TACCTGCTCCTT 2262
```

```
seq_name: /cgn2_6/ptodata/2/1na/backfiles1.seq:5449756-3
```

```
seq_documentation_block:
; Patent No. 5449756
; APPLICANT: TANIGUCHI, TADATSUGU; HATAKEYAMA, MASANORI; MINAMOTO,
; SEIJIRO; KONO, TAKESHI; DOI, TAKESHI; MIYASAKA, MASAYUKI; TSUDO,
; MITSURU; KARASUYAMA, HAJIME
```

```
; TITLE OF INVENTION: RECOMBINANT PROTEIN RECEPTOR FOR IL-2
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/88,592
; FILING DATE: 9-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 865,155
; FILING DATE: 08-APR-1992
; APPLICATION NUMBER: 487,059
; FILING DATE: 05-MAR-1990
; SEQ ID NO:3:
; LENGTH: 2306
5449756-3
```

```
alignment_scores:
  Quality: 74.50      Length: 71
  Ratio: 1.961      Gaps: 5
  Percent Similarity: 53.521      Percent Identity: 29.577
```

```
alignment_block:
US-09-516-493-12 x 5449756-3 ..
```

```
Align seg 1/1 to: 5449756-3 from: 1 to: 2306
```

```
5 ValHisLeu...GlyLeuAlaTrpLeuAla.....GlyP 13
|||||
2074 GTCCACTGCTGAGGAGCTGCTCGCTATCCCTGCCCCCCTCCAAACT 2123
14 .....ValGlySerMetCysLeuPheIleAla.....GlyP 24
2124 TAATCATCCACTTCTGAACCTCATTTGCTACTTCCGTGGCTAACAGGGT 2173
24 heAlaValAlGlyTrpGly.....ProIleProTrp 33
|||||
2174 TTGGTGGAGGGTGGGAGGAGCGGTGCTCAGCTCCACTGCCCTAT 2223
34 LeuLeuMetSerGluIlePheProLeuHisIleYsgIyValAlaThrCl 50
|||||
2224 TTAGTCATG.....AGCTCAGTACGTTCCG 2249
50 yValCysValLeu 54
2250 TACCTGCTCCTT 2262
```

```
seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-781-891-209
```

```
seq_documentation_block:
; Sequence 209, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
```


COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,862
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/701,544
FILING DATE: 16-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6526-049
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1747 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1095
US-08-463-862-1

alignment_scores:
Quality: 70.50 Length: 94
Ratio: 1.905 Gaps: 5
Percent Similarity: 39.362 Percent Identity: 25.532

alignment_block:

US-09-516-493-12 x US-08-463-862-1/rev ..

Align seg 1/1 to reverse of: US-08-463-862-1 from: 1 to: 1747

7 LeuGlyLeuAlaTrpLeuAla..... 13
||||| ||||| |||
1215 CTTGGTCACGCTGCGCTGCCGGGTCCGGTTGAGCAAGTTCATTTC 1166
14ValGlyS 16
1165 TAACATCTCAGAGAGAGACAGTAGACAGACAGCTCCAAAGGAC 1116
16 erMeCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIlePro 32
:: |||||:: |||
1115 AGGGGTCTCTCTCTGTTAGGGGCTC.....TGG.....TGGCCC 1078
33 TrpLeu.....LeuMe 36
|||||
1077 TGGCTGGAAGGGGCTCTCTTGGAGATCAGCTCTTCAGCCGCTCT 1028
36 tSerGluIlePheProLeuHisIleGlyValAlaIleThrGlyValCysV 53
:|||||:: |||||:: |||
1027 TGGGAGATCATCCAGCTCCATGTCAAA.....GGTAATGTC 990
53 alleuThrAsnTrpPheMetaIaPheLeuVal 63
:|||||:: |||||:: |||||
989 TCCTCAGCCACGTGTCATCTGTCGATCATATA 958

seq_name: /cgn2-6/plodata/2/lna/5B_COMB.seq:US-08-461-985-1

seq_documentation_block:

; Sequence 1, Application US/08461985
; Patent No. 5872006
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.

APPLICANT: Yancopoulos, George D.
APPLICANT: Nye, Steven
APPLICANT: Panayiotatos, Nikos
TITLE OF INVENTION: A Family of Map2 Protein Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,985
FILING DATE: 05-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,620
FILING DATE: 03-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6526-123
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1747 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1095
US-08-461-985-1

alignment_scores:
Quality: 70.50 Length: 94
Ratio: 1.905 Gaps: 5
Percent Similarity: 39.362 Percent Identity: 25.532

alignment_block:

US-09-516-493-12 x US-08-461-985-1/rev ..

Align seg 1/1 to reverse of: US-08-461-985-1 from: 1 to: 1747

7 LeuGlyLeuAlaTrpLeuAla..... 13
||||| ||||| |||
1215 CTTGGTCACGCTGCGCTGCCGGGTCCGGTTGAGCAAGTTCATTTC 1166
14ValGlyS 16
1165 TAACATCTCAGAGAGAGACAGTAGACAGACAGCTCCAAAGGAC 1116
16 erMeCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIlePro 32
:|||||:: |||||:: |||
1115 AGGGGTCTCTCTCTGTTAGGGGCTC.....TGG.....TGGCCC 1078
33 TrpLeu.....LeuMe 36
|||||
1077 TGGCTGGAAGGGGCTCTCTTGGAGATCAGCTCTTCAGCCGCTCT 1028
36 tSerGluIlePheProLeuHisIleGlyValAlaIleThrGlyValCysV 53
:|||||:: |||||:: |||


```
1077 TGGCTGGAAGCGGCTGTCTCTTGGAGATCAGCTCTTCAGCGCTCCT 1028
36 tSerGlutIlePheProLeuHisIleLysGlyValAlaThrGlyValCysV 53
   ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
1027 TGGGAGATCATCCAGCTCCATGTCAA.....GGTCAATGGC 990
53 aLeuThrAsnTrpPheMetAlaPheLeuVal 63
   ::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
989 TCCTCAGCCACTGTTCATCTGTGCGATCATA 958
```



```

51 ValCysValLeuThrAsnTrpPheMetAlaPheLeuValThrIysGluPh 67
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
169 GTCTGTTTCTTCCACCAACTGGTTTCATGCGCTTTTGGTGACCAAGAGATT 218
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
67 eAnSerIleMetGluIleLeuArgProTyrgLyAlaPheTrpLeuThra 84
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
219 TAACAGCATATGATGAGATCTTCAGACCCCTAGCGCGCTTGTGCTCACCG 268
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
84 IaAlaPheCysIleLeuSerValLeuPheThr 94
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
269 CTGCTTCTGTATCTCAGCGCTCTTTGCAG 300
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

seq_name: gb_est1:BE231636

seq_documentation_block:
LOCUS BE231636 546 bp mRNA EST 10-JUL-2000
DEFINITION 136437 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BE231636
VERSION BE231636.1 GI:9016354
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
1 (bases 1 to 546)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.,
and Keele,J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
JOURNAL
USDA, ARS, US Meat Animal Research Center
Contact: Smith TPJ
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR PRIMERS
FORWARD: AGGAACAGCTATGACCAAT
BACKWARD: GTTTCAGAGTCACGACG
Plate: 70 row: A column: 21
Seq primer: ATTAGTGACACTATAG.
FEATURES
Location/Qualifiers
source
1..546
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
BASE COUNT
76 a 182 c 162 g 124 t 2 others
ORIGIN
alignment_scores:
Quality: 458.00 Length: 94
Ratio: 4.925 Gaps: 0
Percent Similarity: 98.936 Percent Identity: 86.170

alignment_block:
US-09-516-493-12 x BE231636 ..
Align seg 1/1 to: BE231636 from: 1 to: 546
1 GIUPROAlaSpValIhStleuGlyLeuAlaTrpLeuAlaValIySerMe 17
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2 GAGCCCGCATGCGCGTGGGCTGGCTGGCTGGCTGGCGGCGAGCGT 51
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

```

17 tCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpL 34
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
52 GTGCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 101
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
34 euleMetSerGluIlePheProLeuHsiIleYsgIyAlaIaThrGly 50
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
102 TCTTCATGTCTGAGATCTTCCCTCTGACAGTCAGAGGCGTGGCCACCGGC 151
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
51 ValCysValLeuThrAsnTrpPheMetAlaPheLeuValThrIysGluPh 67
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
152 GTCTGTCTCTCACCAACTGGCTCATGCGCTTCTGTGTGACAAAGAGATT 201
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
67 eAnSerIleMetGluIleLeuArgProTyrgLyAlaPheTrpLeuThra 84
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
202 CAGCAGCCTCATGAGAGTGCTCAGGCGCTTACGCGCTTCTGCTGCGCT 251
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
84 IaAlaPheCysIleLeuSerValLeuPheThr 94
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
252 CGGCTTCTGCATCTCAGTGTCTTTCAC 283
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

seq_name: gb_est2:BG749509

seq_documentation_block:
LOCUS BG749509 949 bp mRNA EST 15-MAY-2000
DEFINITION 602707648F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE 4844411
ACCESSION BG749509
VERSION BG749509.1 GI:14060162
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 949)
NIH-MGC http://mgc.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Contract: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L16M1681 row: P column: 12
High quality sequence stop: 745.
FEATURES
Location/Qualifiers
source
1..949
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:4844411"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV7; Site_1: XhoI, Site
EcoRI; cDNA made by oligo-dT priming. Directly
cloned into EcoRI/XhoI sites using the follo
adaptor: GGCACGAG(G). Library constructed
in the laboratory of Gerald M. Rubin (Univer
California, Berkeley) using ZAP-cDNA synt
(Stratagene) and Superscript II RT (Life
Note: this is a NIH-MGC Library. I"
BASE COUNT
160 a 302 c 264 g 223 t
ORIGIN
alignment_scores:
Quality: 457.00 Length: 94
Ratio: 4.914 Gaps: 0
Percent Similarity: 98.936 Percent Ide

```

alignment_block:
US-09-516-493-12 x BG749509 ..

Align seg 1/1 to: BG749509 from: 1 to: 949

```

1  GluProAlaAspValHisLeuGlyLeuAlaTrpLeuAlaValGlySerMe 17
   ::::::::::::::::::::
80  CAGCGTGTGATGCGAGCGTGGCGCTGGCTGGCGCTGGCGAGCAT 129
   ::::::::::::::::::::
17  tCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProtrPL 34
   ::::::::::::::::::::
130  GTGGCTTCTTCATGCGCGCTTTCGGGCTGGGCGCCATCCCTGGCG 179
   ::::::::::::::::::::
34  euLeuMetSerGluIlePheProLeuHisIleLysGlyValAlaThrGly 50
   ::::::::::::::::::::
180  TCCTCATGTGACGATCTTCCCTCTGCATGTCAAGCGCGTGGCAGACGGC 229
   ::::::::::::::::::::
51  ValCysValLeuThrAsnTrpPheMetAlaPheLeuValThrLysGluPh 67
   ::::::::::::::::::::
230  ATCTCGTCTCTCACCAACTGGCTCATGGCGCTTCTCTGACCAAGAGATT 279
   ::::::::::::::::::::
67  eaSnSerIleMetGluIleLeuArgProTyrGlyAlaPheTrpLeuThra 84
   ::::::::::::::::::::
280  CAGCAGCGCTCATGAGATCTCTCAGCGCTTATGAGAGCGCTTGGCTTGCCT 329
   ::::::::::::::::::::
84  IAlaIlePheCysIleLeuSerValLeuPheThr 94
   ::::::::::::::::::::
330  CCGCTTCTTCATGCTTCATGCTGCTTTCACCT 361

```

seq_name: gb_est2:BF120840

```

seq_documentation_block:
LOCUS      BF120840      823 bp      mRNA      EST      24-OCT-2000
DEFINITION 601757442P1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:3986651 5',
            mRNA sequence.
ACCESSION  BF120840
VERSION    BF120840.1 GI:10959880
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus.

REFERENCE 1 (bases 1 to 823)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LAM9192 row: d column: 12
            High quality sequence start: 4
            High quality sequence stop: 644.

```

FEATURES

```

source
1..823
    location/Qualifiers
    /organism="Mus musculus"
    /strain="C57/B6"
    /db_xref="taxon:10090"
    /clone="IMAGE:3986651"
    /clone_lib="NCI_CGAP_Mam5"
    /tissue_type="tumor, gross tissue"
    /dev_stage="7 months"
    /lab_host="DH10B"
    /note="Organ: mammary; Vector: PCMV-SPORT6; Site: 1; Salt;
    Site: 2; NCI: Cloned unidirectionally. Primer: Oligo dT.
    Library constructed by Life Technologies. Investigators
    providing samples: Lothar Hennighausen/Robin Humphreys,"

```

BASE COUNT 150 a 226 c 257 g 190 t
ORIGIN

alignment_scores:
Quality: 428.50 length: 93
Ratio: 4.761
Percent Similarity: 96.774 Percent Identity: 83.871

alignment_block:
US-09-516-493-12 x BF120840 ..

Align seg 1/1 to: BF120840 from: 1 to: 823

```

1  GluProAlaAspValHisLeuGlyLeuAlaTrpLeuAlaValGlySerMe 17
   ::::::::::::::::::::
204  GAGCTGTGTGATGTCAGAGTGGAGCTGTGCTGGCTGTGAGCAGCAT 253
   ::::::::::::::::::::
17  tCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProtrPL 34
   ::::::::::::::::::::
254  GTGGCTTCTTCATGCTTTCGGGCTGGGCGTGGGAGCCATCCCTGGCG 303
   ::::::::::::::::::::
34  euLeuMetSerGluIlePheProLeuHisIleLysGlyValAlaThrGly 50
   ::::::::::::::::::::
304  TCCATCATGTGACGATCTTCCCTCTGCATGTCAAGCGCTGTGGCTACCGCG 353
   ::::::::::::::::::::
51  ValCysValLeuThrAsnTrpPheMetAlaPheLeuValThrLysGluPh 67
   ::::::::::::::::::::
354  ATCTGTGTCTCTCACCAACTGGCTCATGGCGCTTCTCTGACCAAGAGATT 403
   ::::::::::::::::::::
67  eaSnSerIleMetGluIleLeuArgProTyrGlyAlaPheTrpLeuThra 84
   ::::::::::::::::::::
404  CAGCAGCGCTCATGAGATGCTCAGACCGCTTATGAGAGCGCTTGGCTTGCCT 453
   ::::::::::::::::::::
84  IAlaIlePheCysIleLeuSerValLeuPhe 93
   ::::::::::::::::::::
454  CTGTCTTCTTCTGG...CTCTCAGTGTCTTAT 479

```

seq_name: gb_est2:BE910478

```

seq_documentation_block:
LOCUS      BE910478      934 bp      mRNA      EST      20-OCT-2000
DEFINITION 601501005P1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3903048 5',
            mRNA sequence.
ACCESSION  BE910478
VERSION    BE910478.1 GI:10407108
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens.

REFERENCE 1 (bases 1 to 934)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LAM9706 row: i column: 01
            High quality sequence stop: 710.

```

FEATURES

```

source
1..934
    location/Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:3903048"
    /clone_lib="NIH_MGC_70"

```



```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDF005YE02"
/clone_lib="LRI_FLO13_Fbrn1"
/dev_strage="pooled tissue from post conception fetuses (20
week, 24 week and 26 week)"
/lab_host="DH10B"
/organ="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT      177 a      206 c      242 g      118 t      9 others
ORIGIN

```

```

alignment_scores:
    Quality: 373.50      Length: 98
    Ratio: 4.197      Gaps: 4
    Percent Similarity: 90.816      Percent Identity: 75.510

```

```

alignment_block:
US-09-516-493-12 x AL565390/rev ..

```

```

Align seg 1/1 to reverse of: AL565390 from: 1 to: 752

```

```

1 GluProAlaAspValHisLeuGlyLeuAlaTrpLeuAlaValGlySerme 17
   ::::::::::::::::::::
715 CACGCTGTGATGCGAGCGTGGGCTGGCTGGCTGGCCGTCGNCANCAAT 666
   ::::::::::::::::::::
17 tCysLeuPheIleAla...GlyPheAlaValGlyTrpGlyProIle.Pro 32
   ::::::::::::::::::::
665 GTCCCTCTTTCNATGCGCCGTTTTCGGGTGGGGCCCATCCCT 616
   ::::::::::::::::::::
33 TrpLeuMetSerGluIlePheProLeu.His.IleLysGlyValAla 48
   ::::::::::::::::::::
615 TGCTCTCTATGTCAGAGATCTTCCCTCGCAGCTNTCAAGGGCGTGGC 566
   ::::::::::::::::::::
49 ThrGlyValCysValLeuThrAsnTrpPheMetAlaPheLeuValThrIly 65
   ::::::::::::::::::::
565 ACAGGATGTCGCTCCNACCACTGCGTCATGCGCTTCTCGTAGACCAA 516
   ::::::::::::::::::::
65 scLupheAsnSerIleMetGluIleLeuArgProTyrGlyAlaPheTrpL 82
   ::::::::::::::::::::
515 GGAGTTCACGACGCTCATGAGGTCTCTCAGGCGCTATGAGCCTTGGC 466
   ::::::::::::::::::::
82 eutHraIaIaPheCysIleLeuSerValLeuPheTrpL 94
   ::::::::::::::::::::
465 TTGCTCCGCTTCTGTCATCTTCAGTGTCTCTTTTCACT 428
   ::::::::::::::::::::

```

```

seq_name: gb_est12:BF774206

```

```

seq_documentation_block: 440 bp      mRNA      EST      25-APR-2001
LOCUS      BF774206
DEFINITION 283798 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION  BF774206
VERSION    BF774206.1 GI:12122106
KEYWORDS   EST.
SOURCE     cow.
ORGANISM   Bos taurus

```

```

REFERENCE
1 (bases 1 to 440)
AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
          Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
          ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
          Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and

```

```

TITLE      Keele,J.W.
            Sequence evaluation of four pooled tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
JOURNAL    MEDLINE
GENOME     Res. 11 (4), 626-630 (2001)
COMMENT    Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and a/c trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -mismatch 12 options.
PCR        PCR Primers
FORWARD:   AGGAACAGCTATGACCAT
BACKWARD:  GTTTCCTCCTGACGACG
Plate:     83      row: H      column: 1
Seq primer: ATTAGTGACACTTAG.
            Location/Qualifiers

```

```

FEATURES
    source
        1..440
        /organism="Bos taurus"
        /db_xref="taxon:9913"
        /clone_lib="MARC 3BOV"
        /tissue_type="pooled"
        /lab_host="DH10B"
        /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
        library made from pooled tissue from marrow, alveolar
        macrophage, ovary, fetal semitendinosus muscle, and fetal
        longissimus muscle."

```

```

BASE COUNT      65 a      141 c      135 g      99 t
ORIGIN

```

```

alignment_scores:
    Quality: 351.00      Length: 70
    Ratio: 5.087      Gaps: 0
    Percent Similarity: 98.571      Percent Identity: 90.000

```

```

alignment_block:
US-09-516-493-12 x BF774206 ..

```

```

Align seg 1/1 to: BF774206 from: 1 to: 440

```

```

1 GluProAlaAspValHisLeuGlyLeuAlaTrpLeuAlaValGlySerme 17
   ::::::::::::::::::::
229 GAGGCCGCCGATACCAACGTGGGCTGGCTGGCGCGTGGGCGACGAT 278
   ::::::::::::::::::::
17 tCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpL 34
   ::::::::::::::::::::
279 GTGCTCTTTCATCGCGGCTTCCCTGTGGCTGGGGCCCATCCGTGGC 328
   ::::::::::::::::::::
34 euleuMetSerGluIlePheProLeuHisIleLysGlyValAlaThrGly 50
   ::::::::::::::::::::
329 TCCATGATGTCGAGATCTTCCCTGTCATGTCAGAGGCGTGGCCACCGGC 378
   ::::::::::::::::::::
51 ValCysValLeuThrAsnTrpPheMetAlaPheLeuValThrLysGlyuph 67
   ::::::::::::::::::::
379 GTCTGCTCTCTCACCACTGCTTCATGCGCTTCTGTGTCACCAAGACTT 428
   ::::::::::::::::::::

```

```

67 eAsnSerIle 70
   ::::::::::::::
429 CACGACGTCTC 438
   ::::::::::::::

```

```

seq_name: gb_est1:AM248655

```

```

seq_documentation_block: 496 bp      mRNA      EST      07-JAN-2000
LOCUS      AM248655
DEFINITION 2820759.3prtime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820759 3',
ACCESSION  AM248655
VERSION    AM248655.1 GI:6591648
KEYWORDS   EST.

```


285 CAGCCTGTTGATGCCACCGTGGGCATGGCGCTGCATGG...CCAGTGGG 235

15 yserMeCysLeupheIleAlaGlyPheAlaValGlyTrpGlyProIleP 32
 238 CAGCATGTGCTCTTCATCGCCGCTTTGGCGCTGGCGTGGGGCCCATCC 139

32 roTrpLeuMetSerGluIlePheProLeuHisIleLysGlyValAla 48
 188 CCGGCTCTCATGTCAAGATCTCCCTTCATGTCAAGGGCGTCGCCG 139

49 ThrGlyValCysValLeuThrAsnTrpPheMetAlaPheLeuValThr 65
 138 ACAGCATCTGGCTCCTCACACACATGCTCATGGCTTCTCGTGACACA 89

65 yscIuPheAsnSerIleMetGluIleLeuArgProTyr 77
 88 AGGAGTTCAAGACCGCTCATGAGCTCCTCAGGCCCTAT 51

seq_name: gb_est2:BF140667

seq_documentation_block:
 LOCUS BF140667 917 bp mRNA EST 24-OCT-2000
 DEFINITION 601786917F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4014605 5',
 mRNA sequence.
 ACCESSION BF140667
 VERSION BF140667.1 GI:10979707
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 917)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabos@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM9259 row: a column: 06
 High quality sequence stop: 605.
 Location/Qualifiers
 1..917
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:4014605"
 /clone_lib="NCI_CGAP_Lu30"
 /tissue_type="tumor, metastatic to mammary"
 /lab_host="DH10B"
 /lab="Organ: Lung; Vector: pCMV-SPORT6; Site:1; NCBI:
 Site:2; Salt: transgenic model MMT-1, expression driven by
 MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
 dT. Library constructed by Life Technologies.
 Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 163 a 263 c 277 g 214 t

ORIGIN

alignment_scores:
 Quality: 299.00 Length: 71
 Ratio: 4.333 Gaps: 1
 Percent Similarity: 97.183 Percent Identity: 83.099

alignment_block:
 US-09-516-493-12 x BF140667 ..
 Align seg 1/1 to: BF140667 from: 1 to: 917

1 GluProAlaAspValHisLeuGlyLeuAlaTrpLeuAlaValGlySerMe 17
 430 GAGCCTGTGATGTCACAAAGTGGGACTGGCTGGCTGATGACACAT 479

17 tCySLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProIleP 34
 480 GTCCCTCTTATTTGCTGGCTTTGGGCTGGGACCTATCCCGTGGC 528

34 euleuMetSerGluIlePheProLeuHisIleLysGlyValAlaThrGly 50
 529 T.CTCATGTCAGAGATCTTCCCTTCATGTCAAGGCTGGCTACCGGC 578

51 ValCysValLeuThrAsnTrpPheMetAlaPheLeuValThr.....Ly 65
 579 GTCTGTCTCCTCACACACATGTTTCATGCTCTTTTCTTAAGTTGACCAA 628

65 sgluPheAsnSer 69
 629 AGAGTTCAGAGCA 641

seq_name: gb_est2:BF146289

seq_documentation_block:
 LOCUS BF146289 235 bp mRNA EST 26-OCT-2000
 DEFINITION EST00181 rabbit blastocyst mRNA to cDNA Oryctolagus cuniculus cDNA
 clone G81 similar to glucose transporter 8 (GLUT8), mRNA sequence.
 ACCESSION BF146289
 VERSION BF146289.1 GI:11027684
 KEYWORDS EST.
 SOURCE rabbit.
 ORGANISM Oryctolagus cuniculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 REFERENCE 1 (bases 1 to 235)
 AUTHORS Kietz,S., Augustin,R. and Fischer,B.
 TITLE Expression of glucose transporters in rabbit preimplantation
 embryos
 JOURNAL Unpublished (2001)
 COMMENT Contact: Kietz S
 Department of Anatomy and Cell Biology
 Martin Luther University Halle-Wittenberg, Medical Faculty
 Grosse Steinstrasse 52, D-06097 Halle, GERMANY
 Email: silke.kietz@gmx.de
 Seq primer: 77 Forward
 High quality sequence stop: 235
 POLYA-No.

FEATURES
 source 1..235
 Location/Qualifiers
 1..235
 /organism="Oryctolagus cuniculus"
 /strain="hybrid strain Z1ka"
 /db_xref="taxon:9986"
 /clone="G81"
 /clone_lib="rabbit blastocyst mRNA to cDNA"
 /dev_stage="six days old preimplantation embryo"
 BASE COUNT 39 a 78 c 61 g 57 t

ORIGIN

alignment_scores:
 Quality: 285.00 Length: 61
 Ratio: 4.750 Gaps: 0
 Percent Similarity: 98.361 Percent Identity: 85.246

alignment_block:
 US-09-516-493-12 x BF146289 ..
 Align seg 1/1 to: BF146289 from: 1 to: 235

34 LeuLeuMetSerGluIlePheProLeuHisIleLysGlyValAlaThrGly 50
 2 CTCTCATGTCAGAGATCTTCCGCTGACAGCTCAAGGCTGGCTACGGG 51

```

50 yValcysValleuThrAsnTrpPheMetAlaPheValThrLysGluP 67
|||||
52 CGTCTGGCTCTCACCACACTGGCTCATGGCTTCTGTGATCACCAGAGT 101
|||||
67 heAsnSerIleMetGluIleLeuArgProTyrGlyAlaPheTrpLeuThr 83
|||||
102 TCAGACGCGCATGAGAGCGGCTCGGCGCTTACGCGCGCTTCTGTGCTGCC 151
|||||
84 AlaAlaPheCysTlleLeuSerValleuPheThr 94
|||||
152 TCTGCTCTGTGACATCTTGGGCTCTTTCACCT 184

seq_name: gb_estL2:BG303512

seq_documentation_block:
LOCUS BG303512 295 bp mRNA EST 23-FEB-2001
DEFINITION fl56g08.x1 Sugano Kawakami zebrafish DNA Danio rerio cDNA clone
3817934 3' similar to TR:Q9V610 Q9V610 CG8234 PROTEIN.; mRNA
sequence.
ACCESSION BG303512
VERSION BG303512.1 GI:13101039
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Rasbortinae; Danio.
REFERENCE 1 (bases 1 to 295)
Sugano,S., Kawakami,K., Johnson,S., Li,F., Marra,M., Eddy,S.,
Hillier,L., Clifton,S., Allen,M., Gibbons,M., Jost,S., Kucaba,T.,
Martin,J., Pape,D., Steptoe,M., Underwood,K., Theising,B., Riltter
,E., Bowers,Y., Wylie,T., Waterston,R. and Wilson,R.
Washu Zebrafish EST Project 1999
unpublished (1999)
TITLE JOURNAL
COMMENT Contact: S.L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
Sequencing by: Washington University Genome Sequencing Center Clone
distribution information can be found through the I.M.A.G.E.
Consortium/LNL, send email to: info@image.llnl.gov
Trace considered overall poor quality
Seq primer: 77 from Gibco
High quality sequence stop: 1.
FEATURES
source
location/Qualifiers
1..295
/organism="Danio rerio"
/strain="AB"
/db_xref="taxon:7955"
/clone_lib="Sugano Kawakami zebrafish DNA"
/sex="mixed (one male and one female, including
unfertilized eggs)"
/dev_stage="adult"
/lab_host="DH10B (phage resistant)"
/note="Vector: pME18S-FL3; Site_1: DraIII (CACTGTG);
Site_2: DraIII (CACCATGG); 1st strand cDNA was primed
with an oligo(dT) primer (ATGTGGCTTTTCTTTTCTTTT);
double-stranded cDNA was ligated to a DraIII adaptor
(TGTGGCTCTGG), digested and cloned into distinct DraIII
sites of the pME18S-FL3 vector (5' site CACTGTG, 3' site
CACCATGG). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed by Dr. Sumio Sugano
(University of Tokyo Institute of Medical Science) and
kindly donated by Dr. Koichi Kawakami. Custom primers for
sequencing, 5' end primer CTTTGTCTCTTAAGCTGG and 3' end
primer CGACTCTGAGCTCGAGCA."
BASE COUNT 87 a 82 c 72 g 54 t
ORIGIN

```

```

alignment_scores:
Quality: 282.00 Length: 86
Ratio: 3.972 Gaps: 0
Percent Similarity: 82.558 Percent Identity: 55.814

alignment_block:
US-09-516-493-12 x BG303512/rev ..

Align seq 1/1 to reverse of: BG303512 from: 1 to: 295

9 LeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAl 25
|||||
289 CTGGCTGGTGTCTGTGATAGCATGGCTTTTTCATTCAGAGATTTCG 240
|||||
25 aValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhePro 42
|||||
239 TATGTGTGGGGTCCCTACTCCGTGGCTGTGATGTACAGATCTCCCA 190
|||||
42 euHsIleLysGlyValAlaThrGlyValcysValleuThrAsnTrpPhe 58
|||||
189 CGCGGCGAGGGGATTTAGCCAGTCTGTGTCTTACTACTGAGACC 140
|||||
59 MetAlaPheLeuValThrLysGluPheAsnSerIleMetGluIleLeuAr 75
|||||
139 TGTGCTCTCATTCAGACCAAAACCTTCACACCTCATGTGATGCTGTGAG 90
|||||
75 gProTyrGlyAlaPheTrpLeuThrAlaAlaPheCysTlleLeuSerVal 92
|||||
89 CAGCGCAGGAACATTTGGATGTTCTGCGACAGTGTGCTTAATATGTGC 40
|||||
92 eupHerthr 94
|||||
39 TCTTCAAG 32

seq_name: gb_estL2:BG079217

seq_documentation_block:
LOCUS BG079217 595 bp mRNA EST 26-JAN-2001
DEFINITION H3037C03-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
H3037C03 5', mRNA sequence.
ACCESSION BG079217
VERSION BG079217.1 GI:12561785
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 595)
Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
,T.S., Carter,M.G. and Ko,M.S.H.
Verification and initial annotation of NIA mouse 15K cDNA clone set
unpublished (2001)
OTHER ESTs: H3037C03-3
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6620, USA
Email: cdna@igsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://igsun.grc.nia.nih.gov/cDNA/15k.html for details.
Plate: H3037 row: C column: 03
Seq primer: -21M13 Reverse
High quality sequence stop: 595
POLYA=No.
FEATURES
source
location/Qualifiers
1..595
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="NIA Mouse 15K cDNA Clone Set"

```

/sex="Clones arrayed from a variety of cDNA libraries"
 /dev_stage="Clones arrayed from a variety of cDNA
 libraries"
 /lab_host="DH10B"
 /note="Vector: pSPORT1. Site 1: SalI; Site 2: NotI. This
 clone is among a rearranged set of 15,247 clones from 11
 embryo cDNA libraries (including preimplantation stage
 embryos from unfertilized egg to blastocyst, embryonic
 part of E7.5 embryos, extraembryonic part of E7.5 embryos
 , and E12.5 female mesonephros/gonad) and one newborn
 ovary cDNA library. Average insert size 1.5 kb. All
 source libraries are cloned unidirectionally with Oligo(dT
)-Not primers. References include: (1) Genome-wide
 expression profiling of mid-gestation placenta and embryo
 using a 15,000 mouse developmental cDNA microarray, 2000,
 Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)
 Large-scale cDNA analysis reveals phased gene expression
 patterns during preimplantation mouse development, 2000,
 Development, 127: 1737-1749; (3) Genome-wide mapping of
 unselected transcripts from extraembryonic tissue of
 7.5-day mouse embryos reveals enrichment in the t-complex
 and under-representation on the X chromosome, 1998, Hum
 Mol Genet 7: 1967-1978."

BASE COUNT 95 a 177 c 182 g 141 t
 ORIGIN

alignment_scores:
 Quality: 270.00 Length: 52
 Ratio: 5.192 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 90.385

alignment_block:
 US-09-516-493-12 x BG079217 ..

Align seg 1/1 to: BG079217 from: 1 to: 595

```

1 GIUProAlaSPyAlHISLeuGlyLeuAlATrPLeuAlaValGlySerMe 17
  |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
438 GAGCGCTGGATGTCACAGTGGAGCGCTGGCTGGCTGTAGGACACAT 487
17 TCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpL 34
  |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
488 GTGCCCTTCATTGCTGGCTTTCGGGTGGCGGAGCCATCCCTCCGCG 537
34 euleuMetSerGluIlePheProLeuHISLeuGlyValAlaThrGly 50
  |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
538 TCCTCATGTGACAGATCTTCCCTCTGCGATGTCAAGGCTGTGCGCTACCGCG 587

```

seq_name: gb_est2:BG672321

seq_documentation_block:

LOCUS BG672321 538 bp mRNA EST 30-APR-2001
 DEFINITION DRNCPD10 Rat DRG library Rattus norvegicus cDNA clone DRNCPD10 5',
 mRNA sequence.

ACCESSION BG672321
 VERSION BG672321.1 GI:13894420
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 538)

AUTHORS Xiao,H.S., Han,Z.G., Zhang,F.X., Huang,Q.H., Lu,Y.J., Bao,L., Fu,G.,
 Guo,C., Yan,Q., Jin,S.X., Zhu,Z.D., Xu,X.R., Li,N.G., Chen,Z., and
 Zhang,X.
 TITLE Distinct gene expression profiles of rat dorsal root ganglion
 induced by peripheral nerve axotomy

JOURNAL COMMENT

Unpublished (2001)
 Contact: Zhang Xu
 Laboratory of Sensory System
 Institute of Neuroscience
 320 Yue Yang Road, Shanghai 200031, P.R.China
 Tel: 86-21-64748700-121
 Fax: 86-21-64713446
 Email: xu.zhang@ion.ac.cn
 This clone is also available at Chinese National Human Genome
 Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park,
 Pudong New Area, P.R.China. Please contact with Zhang Xu
 (xu.zhang@ion.ac.cn) or Han Zeguang (hanzeg@chc.sh.cn)
 PCR Primers
 FORWARD: T7
 BACKWARD: T7
 Seq primer: T3
 POLYA-No.

FEATURES

source location/Qualifiers
 1..538
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone_lib="Rat DRG Library"
 /sex="male"
 /tissue_type="dorsal root ganglion"
 /dev_stage="adult"

BASE COUNT 118 a 146 c 138 g 136 t
 ORIGIN

alignment_scores:
 Quality: 269.00 Length: 66
 Ratio: 4.559 Gaps: 0
 Percent Similarity: 89.394 Percent Identity: 77.273

alignment_block:
 US-09-516-493-12 x BG672321 ..

Align seg 1/1 to: BG672321 from: 1 to: 538

```

29 GlyProIleProTrpIleuMetSerGluIlePheProLeuHISLeu 45
  |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
5 GGACCCATTCCTGGCTTCTTATTCAGAGAACCTTCTTGGCACAATAA 54
45 sGlyValAlaThrGlyValCysValLeuThrAsnTrpPheMetAlaPhe 62
  |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
55 GGGGGGGGTACCGGGCTGTGGCTGTCCACCACTGCTTCATGGGCTTTC 104
62 euValThrLySGluPheAsnSerIleMetGluIleLeuArgProTyGly 78
  |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
105 TGGGGACCAAGAGTTTACAGCATCATGGAGATCTTAAACCTACGGC 154
79 AlaPheTrpLeuThrAlaAlaPheCysIleLeuSerValLeuPheThr 94
  |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
155 GCGTTTGGCTCAACGGCTGTCTTTGTATCTGTAGCGGCTTTTCACG 202

```

Thu Feb 14 07:44:32 2002

us-09-516-493-12.rst

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 18:47:17 ; Search time 2841.05 seconds
(without alignments)
3922.273 Million cell updates/sec

Title: US-09-516-493-9

Perfect score: 1037
Sequence: 1 tggcgagcgcctctagaacta.....aaaactcgagggggcc 1037

Scoring table: OLIGO-MDC
Gapop 60.0 , Gapext 60.0

Searched: 11351937 seqs, 5372889281 residues

Word size : 0

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estro:*
9: em_hic:*
10: gb_est1:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_dln:*
18: em_gss_pro:*
19: em_gss_fod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	217	20.9	319	11	H34451 EST111406 R
2	161	15.5	426	11	BF551290 UI-R-CO-h
3	161	15.5	450	11	BF420038 UI-R-BJ2-
4	161	15.5	481	11	AA997295 UI-R-CO-h
5	161	15.5	563	11	BI275901 UI-R-CX0-
6	91	8.8	295	11	BF420257 UI-R-BJ2-
7	70	6.8	324	11	H34372 EST111272 R
8	66	6.4	727	10	AV726694 AV726694
9	66	6.2	723	10	AV706527 AV706527
10	63	6.1	548	10	AI649735 496008612
11	62	6.0	154	10	AW333608 S23H3 AGS
12	62	6.0	234	10	AI894852 EST264295

C 13	62	6.0	235	10	AI649856	AI649856 496007F01
C 14	62	6.0	248	10	AW333091	AW333091 S17B9 AGS
C 15	62	6.0	257	11	BF281142	BF281142 EST1445817
C 16	62	6.0	292	11	BC939722	BC939722 cr61b08..x
C 17	62	6.0	299	10	AW950443	AW950443 EST362513
C 18	62	6.0	305	11	BF019248	BF019248 uy20h9.y
C 19	62	6.0	328	10	AW952183	AW952183 EST364253
C 20	62	6.0	336	10	AW656903	AW656903 AV656903
C 21	62	6.0	339	10	AW955613	AW955613 EST367683
C 22	62	6.0	341	10	AW661704	AW661704 AV661704
C 23	62	6.0	345	10	AV696106	AV696106 AV696106
C 24	62	6.0	345	10	AV697196	AV697196 AV697196
C 25	62	6.0	347	10	AW956075	AW956075 EST368145
C 26	62	6.0	349	10	AV709587	AV709587 AV709587
C 27	62	6.0	363	10	AV650315	AV650315 AV650315
C 28	62	6.0	365	10	AV659389	AV659389 AV659389
C 29	62	6.0	365	10	AW951263	AW951263 EST363333
C 30	62	6.0	366	10	AV727613	AV727613 AV727613
C 31	62	6.0	371	10	AU089893	AU089893 A0089893
C 32	62	6.0	371	10	AV689111	AV689111 AV689111
C 33	62	6.0	371	10	AW333598	AW333598 S23G3 AGS
C 34	62	6.0	373	10	AV655280	AV655280 AV655280
C 35	62	6.0	373	10	AV728157	AV728157 AV728157
C 36	62	6.0	379	10	AW956474	AW956474 EST368544
C 37	62	6.0	380	10	AV659294	AV659294 AV659294
C 38	62	6.0	380	10	AV727787	AV727787 AV727787
C 39	62	6.0	382	10	AV686060	AV686060 AV686060
C 40	62	6.0	382	10	AV692345	AV692345 AV692345
C 41	62	6.0	382	10	AV725745	AV725745 AV725745
C 42	62	6.0	383	10	AV660608	AV660608 AV660608
C 43	62	6.0	383	10	AV686064	AV686064 AV686064
C 44	62	6.0	383	10	AW951239	AW951239 EST363309
C 45	62	6.0	385	10	AV659322	AV659322 AV659322
C 46	62	6.0	385	10	AV726156	AV726156 AV726156
C 47	62	6.0	386	10	AV654908	AV654908 AV654908
C 48	62	6.0	387	10	AV726590	AV726590 AV726590
C 49	62	6.0	388	10	AV656478	AV656478 AV656478
C 50	62	6.0	388	10	AV698545	AV698545 AV698545
C 51	62	6.0	394	10	AV708893	AV708893 AV708893
C 52	62	6.0	394	10	AV709314	AV709314 AV709314
C 53	62	6.0	395	10	AV653353	AV653353 AV653353
C 54	62	6.0	396	10	AV708381	AV708381 AV708381
C 55	62	6.0	398	10	AV660728	AV660728 AV660728
C 56	62	6.0	402	10	AV659536	AV659536 AV659536
C 57	62	6.0	404	10	AV691080	AV691080 AV691080
C 58	62	6.0	406	10	AV702385	AV702385 AV702385
C 59	62	6.0	406	10	AV706219	AV706219 AV706219
C 60	62	6.0	411	10	AV652001	AV652001 AV652001
C 61	62	6.0	411	10	AV651955	AV651955 AV651955
C 62	62	6.0	417	10	AV707979	AV707979 AV707979
C 63	62	6.0	421	10	AV709580	AV709580 AV709580
C 64	62	6.0	421	10	AV727003	AV727003 AV727003
C 65	62	6.0	430	10	AV708786	AV708786 AV708786
C 66	62	6.0	431	10	AV659547	AV659547 AV659547
C 67	62	6.0	437	10	AV648263	AV648263 AV648263
C 68	62	6.0	437	10	AV727526	AV727526 AV727526
C 69	62	6.0	439	10	AV703169	AV703169 AV703169
C 70	62	6.0	454	10	AV728518	AV728518 AV728518
C 71	62	6.0	456	10	AV707541	AV707541 AV707541
C 72	62	6.0	465	10	AV725033	AV725033 AV725033
C 73	62	6.0	468	10	AV728924	AV728924 AV728924
C 74	62	6.0	469	10	AV725617	AV725617 AV725617
C 75	62	6.0	485	10	AV696931	AV696931 AV696931
C 76	62	6.0	502	10	AV699089	AV699089 AV699089
C 77	62	6.0	505	10	AV705135	AV705135 AV705135
C 78	62	6.0	508	10	AV701874	AV701874 AV701874
C 79	62	6.0	512	10	AV704637	AV704637 AV704637
C 80	62	6.0	516	10	AW333040	AW333040 S16E5 AGS
C 81	62	6.0	529	10	AV703501	AV703501 AV703501
C 82	62	6.0	551	10	AA753446	AA753446 96BS0473
C 83	62	6.0	556	10	AW331977	AW331977 S1D4 AGS-
C 84	62	6.0	561	10	AV701180	AV701180 AV701180
C 85	62	6.0	567	10	AI649814	AI649814 496006E10

SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 481)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT On Jun 5, 1998 this sequence version replaced g1:3188156.
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized adult 12-day-embryo library. CDNA Library Preparation: M. Fatima Bonaldo, Ph.D. Clone distribution: clones will be available through Research Genetics This clone is also available through the I.M.A.G.E. Consortium at LBNL (info@image.lbnl.gov). IMAGE ID=1773689
Seq primer: M13 Forward
POLYA=NO.

FEATURES
Source
Location/Qualifiers
1..481
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C0-bq-f-01-0-UI"
/clone_lib="UI-R-C0"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified polylinker. Site.1: Not I; Site.2: Eco RI; The UI-R-C0 library is a subtracted library derived from the UI-R-A1 and UI-R-E1 libraries. The UI-R-A1 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, and muscle. The UI-R-E1 library consisted of a mixture of individually tagged normalized libraries constructed from 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C0) was constructed as follows: PCR amplified cDNA inserts from a pool of UI-R-A1 and UI-R-E1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the pooled UI-R-A1 and UI-R-E1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C0 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)".

BASE COUNT 118 a 111 c 128 g 124 t
ORIGIN

Query Match 15.5%; Score 161; DB 10; Length 481;
Best Local Similarity 100.0%; Pred. No. 1.1e-27;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 634 tgcgggctccctggtctagctctgctgctgacatcttggtgggtgctgtgctcaagcaa 693
Db 364 tgcgggctccctggtctagctctgctgctgacatcttggtgggtgctgtgctcaagcaa 305

Qy 694 ctgccatctactctactgtgctggggatagagaaggacttagccacataagattgggc 753
Db 304 ctgccatctactctactgtgctggggatagagaaggacttagccacataagattgggc 245

Qy 754 tcagaacaaggctcaggtagctccaggaagaagaagatg 794
Db 244 tcagaacaaggctcaggtagctccaggaagaagaagatg 204

RESULT 5
BI275901/c
LOCUS BI275901 563 bp mRNA EST 19-JUL-2001
DEFINITION UI-R-CX0-bxh-h-03-0-UI s1 UI-R-CX0 Rattus norvegicus cDNA clone
ACCESSION UI-R-CX0-bxh-h-03-0-UI 3', mRNA sequence.
VERSION BI275901
KEYWORDS BI275901.1 GI:14920290
SOURCE EST.
ORGANISM Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 563)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized rat placenta pool library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=yes.

FEATURES
Source
Location/Qualifiers
1..563
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CX0-bxh-h-03-0-UI"
/clone_lib="UI-R-CX0"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified polylinker. Site.1: Not I; Site.2: Eco RI; The UI-R-CX0 library is a normalized library constructed from the following rat placenta tissues: embryonic day 17, embryonic day 19, embryonic day 21. For a detailed description of the library from which this clone was derived, please visit our web site at ratseq.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)
TAG_LIB=UI-R-CX0
TAG_TISSUE=rat placenta pool
TAG_SEQ=TCACGACAGT"

BASE COUNT 135 a 141 c 163 g 124 t
ORIGIN

Query Match 15.5%; Score 161; DB 11; Length 563;

Best Local Similarity 100.0%; Pred. No. 9.7e-28;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 634 tgcggagctcctgcagtcgtcgtgcatcttctggcgtgcgtcccaaca 693
|||||
Db 310 tgcggagctcctgcagtcgtcgtgcatcttctggcgtgcgtcccaaca 251
|||||
Oy 694 ctgcacatacctcactgactgagtgagaaaggactagcacataaattgggc 753
|||||
Db 250 ctgcacatacctcactgactgagtgagaaaggactagcacataaattgggc 191
|||||
Oy 754 tcgaaacaagtcagtcagtcagtcaggaagaagaagaatg 794
|||||
Db 190 tcgaaacaagtcagtcagtcagtcaggaagaagaagaatg 150
|||||

RESULT 6

BF420257/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

1 (bases 1 to 295)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized ventricle at 16.5 dpc library cDNA Library Preparation:
M.B. Soares Lab Clone distribution: clones will be available
through Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 1-21,
>AT rich; low complexity
Seq primer: M13 Forward
POLYA=yes.

FEATURES
source
Location/Qualifiers
1..295

/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="01-R-BJ2-bpk-g-03-0-01"
/lab_host="DH10B (Life Technologies)"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The 01-R-BJ2
library is a subtracted library derived from the following
tissues: heart, atrium at 15 dpc, ventricle at 16.5 dpc,
atrium at 16.5 dpc, ventricle at 13 dpc, ventricle at 15
dpc, AV canal at 15 dpc. For a detailed description of
the library from which this clone was derived, please
visit our web site at ratel.eng.uiowa.edu. The
subtraction has been previously described in (Bonaldo,
Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_L1B=01-R-BJ2

TAG_TISSUE=ventricle at 16.5 dpc
TAG_SEQ=GTTCG
BASE COUNT 76 a 63 c 68 g 88 t
ORIGIN

Query Match 8.8%; Score 91; DB 11; Length 295;
Best Local Similarity 100.0%; Pred. No. 9.8e-12;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 704 ctccactgactgggggagtgagaaaggactagcacataaattggcgcagaaca 763
|||||
Db 295 ctccactgactgggggagtgagaaaggactagcacataaattggcgcagaaca 236
|||||
Oy 764 gtcagtgagtcaggaagaagaagaatg 794
|||||
Db 235 gtcagtgagtcaggaagaagaagaatg 205
|||||

RESULT 7

H34372/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

1 (bases 1 to 324)
Lee, N.H., Weinstein, K.G., Kirkness, E.F., Earle-Hughes, J.A., Fuldner
, R.A., Marmaras, S., Glodok, A., Gocayne, J.D., Adams, M.D., Kerlavage
, A.R., Fraser, C.M., and Venter, J.C.
Comparative expressed-sequence-tag analysis of differential gene
expression profiles in PC-12 cells before and after nerve growth
factor treatment
Proc. Natl. Acad. Sci. U.S.A. 92, 8303-8307 (1995)
95396786
Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
For clone availability please contact the TIGR Database
(tdbinfo@tdb.tigr.org)
Seq primer: M13 - 21.
Location/Qualifiers
1..324

FEATURES
source

/organism="Rattus sp."
/db_xref="ATCC (inhost):2005377"
/db_xref="taxon:10118"
/clone="RPMBT28"
/clone_lib="Rat PC-12 cells, NGF-treated (9 days)"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; poly(A)+ RNA was purified from 9-day NGF treated
PC12 cells. cDNA was constructed using an oligo-dT primer
and directionally cloned using the Lambda Zap II Vector
Kit by Stratagene"
BASE COUNT 89 a 74 c 79 g 78 t 4 others
ORIGIN

Query Match 6.8%; Score 70; DB 11; Length 324;
Best Local Similarity 100.0%; Pred. No. 4.8e-07;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 725 aaaggacttgcacataaattggcgtcagaaacaggtcaggtagtcaggaaga 784
|||||
Db 256 aaaggacttgcacataaattggcgtcagaaacaggtcaggtagtcaggaaga 197
|||||

QY 785 aaagaatg 794
|||||
Db 196 AAAGAGATG 187

RESULT 8
AV726694
LOCUS AV726694 727 bp mRNA EST 17-OCT-2000
DEFINITION AV726694 HTC Homo sapiens cDNA clone HTCBUH09 5', mRNA sequence.
ACCESSION AV726694
VERSION AV726694.1 GI:10836115
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 727)
AUTHORS Li N., Qian B., Song Y., Song H., Huang Q., Yang Y., Gao G., Xiao H., Xu X., Gu Y., Peng Y., Liu F., Qu J., Gao X., Cheng Z., Xu Z., Zeng L., Xu S., Gu W., Tu X., Jia J., Fu G., Ren S., Zhong M., Lu G., Hu R., Chen J., Chen Z. and Han Z.
Homo sapiens cDNA HTC clones

TITLE Unpublished (2000)
JOURNAL Contact: Zequang Han
COMMENT Chinese National Human Genome Center at Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzgchc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
SOURCE
1..727
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HTCUBH09"
/clone_lib="HTC"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 248 a 150 c 164 g 163 t 2 others
ORIGIN

Query Match 6.4%; Score 66; DB 10; Length 727;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 972 gtctcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagg 1031
|||||
Db 370 GTTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATCTGAGGG 429

QY 1032 gggggc 1037
|||||
Db 430 GGGGCC 435

RESULT 9
AV706527
LOCUS AV706527 723 bp mRNA EST 09-OCT-2000
DEFINITION AV706527 ADB Homo sapiens cDNA clone ADBCMF12 5', mRNA sequence.
ACCESSION AV706527
VERSION AV706527.1 GI:10723806
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 723)
AUTHORS Peng Y., Song H., Huang Q., Huang C., Gu Y., Yang Y., Gao G., Xiao

H., Xu X., Li N., Qian B., Liu F., Qu J., Gao X., Cheng Z., Xu Z., Zeng L., Xu S., Gu W., Tu Y., Jia J., Fu G., Ren S., Zhong M., Lu G., Hu R., Chen J., Chen Z. and Han Z.
Homo sapiens cDNA ADB clones
Unpublished (2000)
JOURNAL Contact: Zequang Han
COMMENT Chinese National Human Genome Center at Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzgchc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
SOURCE
1..723
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ADBCMF12"
/clone_lib="ADB"
/tissue_type="Adrenal gland"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 225 a 158 c 149 g 191 t
ORIGIN

Query Match 6.2%; Score 64; DB 10; Length 723;
Best Local Similarity 100.0%; Pred. No. 5.1e-06;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 974 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagg 1033
|||||
Db 279 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATCTGAGGG 338

QY 1034 gggc 1037
|||||
Db 339 GGCC 342

RESULT 10
A1649735
LOCUS A1649735 548 bp mRNA EST 30-APR-1999
DEFINITION 496008G12.x1 496 - stressed shoot cDNA library from Wang/Bohnett
1ad Zea mays cDNA, mRNA sequence.
ACCESSION A1649735
VERSION A1649735.1 GI:4730569
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays

REFERENCE 1 (bases 1 to 548)
AUTHORS Walbot V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University

JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 496008 row: G column: 12.

FEATURES
SOURCE
1..548
Location/Qualifiers
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="496 - stressed shoot cDNA library from

```

Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E.coli XL Gold"
/note="Organ: shoot; Vector: pBluescriptII SK(+); XR:
Wang/Bohnert"

BASE COUNT      157 a      123 c      113 g      155 t
ORIGIN

Query Match      6.1%: Score 63; DB 10; Length 548;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 975 caaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgaggagg 1034
|||||
Db 66 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTCGAGGGGG 7

QY 1035 gcc 1037
|||
Db 6 gcc 4

RESULT 11
AM333608      154 bp      mRNA      EST      31-JAN-2000
LOCUS      523H3 AGS-1 Pneumocystis carinii f. sp. carinii cDNA 3', mRNA
DEFINITION
sequence.
ACCESSION      AM333608.1 GI:6829965
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Pneumocystis carinii f. sp. carinii.
Pneumocystis carinii f. sp. carinii
Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;
Pneumocystidaceae; Pneumocystis.
1 (bases 1 to 154)
Smullan,A.G., Arnold,J., Weise,M., Wunderlich,J., Staben,C., Edman
,J.C., Kovacs,J. and Cushion,M.
Expressed sequence tags from Pneumocystis carinii
Unpublished (2000)
Contact: Staben C
School of Biological Sciences
University of Kentucky
101 Morgan Building, University of Kentucky, Lexington, KY
40506-0225, USA
Tel: 606 257 2161
Fax: 606 257 1717
Email: staben@pop.uky.edu.
Location/Qualifiers
1. 154
/organism="Pneumocystis carinii f. sp. carinii"
/db_xref="taxon:38081"
/clone_lib="AGS-1"
/lab_host="E. coli"
/note="Vector: Lambda ZAP II; Site_1: EcoRI; Site_2: XhoI;
P. carinii organisms (3x10e9) from a single rat (99-1-6,
sacrificed on 3/17/99) at Cincinnati VA facilities.
Tritol extracted RNA. Oligo dT priming, standard
conditions described by vendor, Stratagene. Further
details see www.uky.edu/Project/Pneumocystis/"

BASE COUNT      36 a      14 c      14 g      90 t
ORIGIN

Query Match      6.0%: Score 62; DB 10; Length 154;
Best Local Similarity 100.0%; Pred. No. 5.9e-05;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgaggagg 1035
|||||
Db 66 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTCAGGGGGG 7

QY 1036 cc 1037

```

```

Db 6 cc 5

RESULT 12
A1894852
LOCUS      A1894852      234 bp      mRNA      EST      18-MAY-2001
DEFINITION      EST264295 tomato callus, TAMU Lycopersicon esculentum cDNA clone
CLEC6G13, mRNA sequence.
ACCESSION      A1894852
VERSION
KEYWORDS
EST.
SOURCE      Lycopersicon esculentum
ORGANISM      Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 234)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S., Ronning
,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato callus tissue
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
Location/Qualifiers
1. 234
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone_lib="CLEC613"
/clone_lib="tomato callus, TAMU"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="XLI-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Giovannoni Laboratory; cLEC - cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
Mixed callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST Library"

BASE COUNT      105 a      35 c      31 g      63 t
ORIGIN

Query Match      6.0%: Score 62; DB 10; Length 234;
Best Local Similarity 100.0%; Pred. No. 4e-05;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgaggagg 1035
|||||
Db 153 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTCAGGGGGG 212

QY 1036 cc 1037
|||
Db 213 CC 214

RESULT 13
A1649856
LOCUS      A1649856      235 bp      mRNA      EST      30-APR-1999
DEFINITION      496007F01.x1 496 - stressed shoot cDNA library from Wang/Bohnert
lab Zea mays cDNA, mRNA sequence.
ACCESSION      A1649856
VERSION      A1649856.1 GI:4730690
KEYWORDS
EST.
SOURCE      Zea mays.
ORGANISM      Zea mays.

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 235)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Maize ESTs from various cDNA libraries sequenced at Stanford University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 496007 row: F column: 01.
Location/Qualifiers

FEATURES

source
1..235
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="496 - stressed shoot cDNA library from Wang/Bohner lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E.coli XL Gold"
/note="Organ: shoot; Vector: pBluescriptII SK(+) XR; Wang/Bohner"

BASE COUNT

55 a 48 c 50 g 82 t

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 62; DB 10; Length 235;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagggggg 1035

Db 68 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCTCAGGGGGG 9

OY 1036 cc 1037

||

8 CC 7

RESULT 14

AM333091 248 bp mRNA EST 31-JAN-2000
LOCUS S17B9 AGS-1 Pneumocystis carinii f. sp. carinii cDNA 3', mRNA
DEFINITION
Sequence.

ACCESSION

AM333091
AM333091.1 GI:6829448

VERSION

EST.

KEYWORDS

Pneumocystis carinii f. sp. carinii.

SOURCE

Pneumocystis carinii f. sp. carinii.
Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;
Pneumocystidaceae; Pneumocystis.

ORGANISM

Smilian, A.G., Arnold, J., Weise, M., Wunderlich, J., Staben, C., Edman, J.C., Kovacs, J. and Cushion, M.
Expressed sequence tags from Pneumocystis carinii
Unpublished (2000)

REFERENCE

1 (bases 1 to 248)

AUTHORS

Smilian, A.G., Arnold, J., Weise, M., Wunderlich, J., Staben, C., Edman, J.C., Kovacs, J. and Cushion, M.

TITLE

Expressed sequence tags from Pneumocystis carinii

JOURNAL

Unpublished (2000)

COMMENT

Contact: Staben C
School of Biological Sciences
University of Kentucky
101 Morgan Building, University of Kentucky, Lexington, KY
40506-0225, USA
Tel: 606 257 2161
Fax: 606 257 1717
Email: staben@pop.uky.edu.
Location/Qualifiers

FEATURES

source

FEATURES

source

/organism="Pneumocystis carinii f. sp. carinii"

/db_xref="taxon:38081"
/clone_lib="AGS-1"
/lab_host="E. coli"
/note="Vector: Lambda ZAP II; Site_1: EcoRI; Site_2: XhoI;
P. carinii organisms (3x10e9) from a single rat (99-1-6,
sacrificed on 3/17/99) at Cincinnati VA facilities.
Trizol extracted RNA. Oligo dt priming, standard
conditions described by vendor. Stragene. Further
details see www.uky.edu/project/pneumocystis/"

BASE COUNT

37 a 38 c 54 g 119 t

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 62; DB 10; Length 248;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagggggg 1035

Db 65 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCTCAGGGGGG 6

OY 1036 cc 1037

||

5 CC 4

RESULT 15

BF281142 257 bp mRNA EST 28-NOV-2000
LOCUS EST45817 Rat Gene Index, normalized rat, Rattus norvegicus cDNA
DEFINITION
Rattus norvegicus cDNA clone RGIAB64 3' sequence, mRNA sequence.
BF281142
BF281142.1 GI:11212296

ACCESSION

EST.

VERSION

EST.

KEYWORDS

EST.

SOURCE

EST.

ORGANISM

Rattus norvegicus

REFERENCE

1 (bases 1 to 257)

AUTHORS

Malek, R.L., Cho, J., Lee, Y., Karamycheva, S., Parvizi, B., Pertea, G.,
Sultana, R., Tsai, J., White, J., Quackenbush, J. and Lee, N.H.
Generation of ESTs from Normalized Rat Embryo, Bento Soares
Unpublished (2000)

TITLE

Other ESTs: EST345498

JOURNAL

Contact: Lee, NH

COMMENT

The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
This clone is available through the ATCC, contact the ATCC
tel#703-365-2700 for further information.

FEATURES

source

FEATURES

source

FEATURES

source

BASE COUNT

58 a 38 c 53 g 108 t

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 62; DB 11; Length 257;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -
 XX
 XX Claim 11: Page 71-72; 124pp; English.
 CC The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is rat GLUTX1.
 CC
 SQ Sequence 478 AA:
 Query Match 70.5%; Score 627; DB 22; Length 478;
 Best Local Similarity 100.0%; Pred. No. 3.1e-64;
 Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 19 LVPISEPADVHLGLAMLVAGSMCLFIAGFVAGWGPIPWLMSEIFPLHIKGVATGCVL 78
 DB 357 lvpisaeppdvhlglawlgvsmclfiagfavgwgpipwlmselplhikgvatgcvl 416
 QY 79 TWMFMAFLVTKKEFNSIMEILRPYGAFWLTAFCLISVLTFTLFVPEKGRTELOITAH 136
 DB 417 twmfmaflvtkefnsimeilrpygafwltafclisvltftlfpvpekgrtleqitah 474

RESULT 2
 ID AAB66939 standard; Protein; 478 AA.
 XX
 AC AAB66939;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE GLUTX1 consensus sequence.
 XX
 KM GLUTX: gene therapy: vaccine: hexose transport modulator; human; rat;
 KM hexose transport disorder: ischaemia; diabetes: hyperglycaemia; murine;
 KM hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 OS Homo sapiens.
 OS Mus sp.
 OS Rattus sp.
 XX
 PN WO200104145-A2.
 XX
 PD 18-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-IB01042.
 XX
 PR 14-JUL-1999; 99US-0143907.
 PR 27-AUG-1999; 99US-0151140.
 PR 23-FEB-2000; 2000US-0184285.
 PR 13-JUL-2000; 2000US-0616132.
 XX
 PA (UYLA-) UNIV LAUSANNE.
 XX
 PI Thorens B, Ibberson M, Uldry M;
 XX
 DR WPI: 2001-112615/12.
 XX
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -
 XX
 XX Claim 11: Page 74-75; 124pp; English.
 CC The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative

CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is a consensus sequence
 CC for GLUTX1.
 CC
 SQ Sequence 478 AA:
 Query Match 67.3%; Score 598; DB 22; Length 478;
 Best Local Similarity 93.2%; Pred. No. 7.2e-61;
 Matches 110; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 QY 19 LVPISEPADVHLGLAMLVAGSMCLFIAGFVAGWGPIPWLMSEIFPLHIKGVATGCVL 78
 DB 357 lvpisaeppdvhlglawlgvsmclfiagfavgwgpipwlmselplhikgvatgcvl 416
 QY 79 TWMFMAFLVTKKEFNSIMEILRPYGAFWLTAFCLISVLTFTLFVPEKGRTELOITAH 136
 DB 417 twmfmaflvtkefnsimeilrpygafwltafclisvltftlfpvpekgrtleqitah 474

RESULT 3
 ID AAB66934 standard; Protein; 477 AA.
 XX
 AC AAB66934;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE Murine GLUTX1.
 XX
 KM Murine; GLUTX: gene therapy: vaccine: hexose transport modulator;
 KM hexose transport disorder: ischaemia; diabetes: hyperglycaemia;
 KM hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 OS Mus sp.
 OS WO200104145-A2.
 XX
 PD 18-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-IB01042.
 XX
 PR 14-JUL-1999; 99US-0143907.
 PR 27-AUG-1999; 99US-0151140.
 PR 23-FEB-2000; 2000US-0184285.
 PR 13-JUL-2000; 2000US-0616132.
 XX
 PA (UYLA-) UNIV LAUSANNE.
 XX
 PI Thorens B, Ibberson M, Uldry M;
 XX
 DR WPI: 2001-112615/12.
 DR N-PSDB: AAF55867.
 XX
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 PT the prevention, diagnosis and treatment of hexose transport disorders,
 PT e.g. ischemia and diabetes -
 XX
 XX Claim 11: Page 73-74; 124pp; English.
 CC The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is murine GLUTX1.
 CC
 SQ Sequence 477 AA:

PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 the prevention, diagnosis and treatment of hexose transport disorders,
 e.g., ischemia and diabetes -

Claim 11; Page 71-72; 124pp; English.

CC The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and
 CC AAF66932-AA66941). The GLUTX proteins are related to the facilitative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is rat GLUTX1.

SO Sequence 478 AA;

Query Match 100.0%; Score 508; DB 22; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1,36-55;

Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPADVHLGLAWLAVGSMCLFIAGFAVGWGPIMLMSFIPPLIKGVATGCVLTNFMMA 60
 Db 363 epadvhlglawlavgs mclfiagfavgwgpipwlmselfpplhkvatgcvltwfmma 422
 QY 61 FLVTKERNSTIMEILRPYGAFMLTAPFCILSVLFT 94
 Db 423 flvtkernstimeilrpygafmltaafclslvlt 456

RESULT 2

AAB66939
 ID AAB66939 standard; Protein; 478 AA.

AC AAB66939;

DT 17-APR-2001 (first entry)

DE GLUTX1 consensus sequence.

KW GLUTX; gene therapy; vaccine; hexose transport modulator; human; rat;
 KW hexose transport disorder; ischemia; diabetes; hyperglycaemia; murine;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.

OS Homo sapiens.

OS Mus sp.

PN Rattus sp.

PD WO200104145-A2.

PF 14-JUL-2000; 2000WO-IB01042.

PR 14-JUL-1999; 99US-0143907.

PR 27-AUG-1999; 99US-0151140.

PR 23-FEB-2000; 2000US-0184285.

PR 13-JUL-2000; 2000US-0616132.

PA (UYLA-) UNIV LAUSANNE.

PI Thorens B, Ibberson M, Uldry M;

DR WPI; 2001-112615/12.

PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 the prevention, diagnosis and treatment of hexose transport disorders,
 e.g., ischemia and diabetes -

Claim 11; Page 74-75; 124pp; English.

The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and
 AAF66932-AA66941). The GLUTX proteins are related to the facilitative

CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is a consensus sequence
 for GLUTX1.

SO Sequence 478 AA;

Query Match 95.9%; Score 487; DB 22; Length 478;
 Best Local Similarity 92.6%; Pred. No. 5,66-53;

Matches 87; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPADVHLGLAWLAVGSMCLFIAGFAVGWGPIMLMSFIPPLIKGVATGCVLTNFMMA 60
 Db 363 epadvhlglawlavgs mclfiagfavgwgpipwlmselfpplhkvatgcvltwfmma 422
 QY 61 FLVTKERNSTIMEILRPYGAFMLTAPFCILSVLFT 94
 Db 423 flvtkernstimeilrpygafmltaafclslvlt 456

RESULT 3

AAB66934
 ID AAB66934 standard; Protein; 477 AA.

AC AAB66934;

DT 17-APR-2001 (first entry)

DE Murine GLUTX1.

KW Murine; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischemia; diabetes; hyperglycaemia;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.

OS Mus sp.

PN WO200104145-A2.

PF 14-JUL-2000; 2000WO-IB01042.

PR 14-JUL-1999; 99US-0143907.

PR 27-AUG-1999; 99US-0151140.

PR 23-FEB-2000; 2000US-0184285.

PR 13-JUL-2000; 2000US-0616132.

PA (UYLA-) UNIV LAUSANNE.

PI Thorens B, Ibberson M, Uldry M;

DR WPI; 2001-112615/12.

DR N-PSDB; AAF55867.

PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 the prevention, diagnosis and treatment of hexose transport disorders,
 e.g., ischemia and diabetes -

Claim 11; Page 73-74; 124pp; English.

CC The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and
 CC AAF66932-AA66941). The GLUTX proteins are related to the facilitative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is murine GLUTX1.

SO Sequence 477 AA;

Thu Feb 14 07:44:46 2002

Query Match 99.0%; Score 2301.5; DB 22; Length 477;
 Best Local Similarity 99.1%; Pred. No. 3.4e-223;
 Matches 450; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

1 RRVFLAFAALGSLSPFALGYSPPAIPSLQRAAPAPRLDDAAAFGAVVTGGAAG 60
 24 rrvflfaafaalgpstfgfalgysspaipslqraappriiddaaswfgavvllgaaag 83
 61 GVLGGMLVDRAGRKLSLLCSVPFVAGFAVITTAODVWMLLGGRLTGLACGASLVA 120
 84 gvlggwlvdragrkslllcsvfpvagfavltlaeqvwmlllggrrlltglaqvastlvapv 143
 121 YISEIAYPAVAVGLGSCVOLMVVGLILAYLAGVLEWRMLAVLCVPPSLMLLMCFMP 180
 144 yiseiaypavavgligscvqlmvvvglllaylagvlewrlavlgcvppslmllmctmp 203
 181 ETPRELLTQHRROEAI-ALREFLMGSEOGMEDPRIGAEOSFHIALRQPGIYKPFITIGVSL 239
 204 etprflltqhrrgeamaalrfllwgseegwepdpriagsgfnhalltrpglykpflligvsl 263
 240 MAFOOLSGVNAVVFYAEITFEFAKFKDSSLASVVVGVITQVLTAVAAALIMDRAGRLLLV 299
 264 mafoolsgvnavvfyaetifeeakfkdsllasvvvgvltqvltavaalimdragrlllv 323
 300 LSGVVMWFSTSAFGAYFKLTGCGPGNSHVAISAPVSAOPVDASVGLAMLAAGSMCLFIA 359
 324 lsgvwmwfstsaafayfkltdgpgpnsahvasapvsagvdasvglawlagvmmclfia 383
 360 GFAVAGWGPFPWLLMSEIFPLHVKGVANGICVLTNMLMAFLVTKFESSLMELVLRPGAFWL 415
 384 gfavagwgpfpwllmseifplhvkvgangicvltnwlmalfvltkesslmevlrpgafwl 443
 420 ASARCFISVLTFLFCVPERTKGKTLEOITAHFEGR 453
 444 asarcflsvltflfcvpeikgktleqitahfegr 477

RESULT 3
 AAB66939 standard; Protein: 478 AA.
 ID AAB66939 standard; Protein: 478 AA.
 AC AAB66939;
 DT 17-APR-2001 (first entry)
 DE GLUTX1 consensus sequence.
 KW GLUTX: gene therapy; vaccine; hexose transport modulator; human; rat;
 KM hexose transport disorder; ischemia; diabetes; hyperglycaemia; murine;
 hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 OS Homo sapiens.
 OS Mus sp.
 OS Rattus sp.
 PN MO200104145-A2.
 PD 18-JAN-2001.
 PF 14-JUL-2000; 2000MO-1B01042.
 PR 14-JUL-1999; 99US-0143907.
 PR 27-AUG-1999; 99US-0151140.
 PR 23-FEB-2000; 2000US-0184285.
 PR 13-JUL-2000; 2000US-0516132.
 PA (UYLA-) UNIT LAUSANNE.
 PI Thorens B, Ibberson M, Uldry M;
 DR MPI; 2001-112615/12.
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in

the prevention, diagnosis and treatment of hexose transport disorders,
 e.g. ischemia and diabetes -
 Claim 11: Page 74-75; 124pp; English.

The present invention relates to GLUTX proteins (AA055865-AA055871 and
 AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
 glucose carriers GLUT3-GLUT5 and have hexose binding and/or transport
 function. The GLUTX proteins may be used in the diagnosis, prevention and
 treatment of hexose transport disorders such as ischaemia, diabetes,
 hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 neurodegenerative disease. The present sequence is a consensus sequence
 for GLUTX1.

Sequence 478 AA;
 SQ

Query Match 90.5%; Score 2104; DB 22; Length 478;
 Best Local Similarity 89.0%; Pred. No. 2.7e-203;
 Matches 405; Conservative 22; Mismatches 26; Indels 2; Gaps 2;

1 RRVFLAFAALGSLSPFALGYSPPAIPSLQRAAPAPRLDDAAAFGAVVTGGAAG 60
 24 rrvflfaafaalgpstfgfalgysspaipslqraappriiddaaswfgavvllgaaag 83
 61 GVLGGMLVDRAGRKLSLLCSVPFVAGFAVITTAODVWMLLGGRLTGLACGASLVA 120
 84 gvlggwlvdragrkslllcsvfpvagfavltlaeqvwmlllggrrlltglaqvastlvapv 143
 121 YISEIAYPAVAVGLGSCVOLMVVGLILAYLAGVLEWRMLAVLCVPPSLMLLMCFMP 180
 144 yiseiaypavavgligscvqlmvvvglllaylagvlewrlavlgcvppslmllmctmp 203
 181 ETPRELLTQHRROEAI-ALREFLMGSEOGMEDPRIGAE-OSFHIALRQPGIYKPFITIGVSL 238
 204 etprflltqhrrgeamaalrfllwgseegwepdpriagsgfnhalltrpglykpflligvsl 263
 239 LMAFOOLSGVNAVVFYAEITFEFAKFKDSSLASVVVGVITQVLTAVAAALIMDRAGRLLLV 298
 264 lmafoolsgvnavvfyaetifeeakfkdsllasvvvgvltqvltavaalimdragrlllv 323
 299 VLSGVVMWFSTSAFGAYFKLTGCGPGNSHVAISAPVSAOPVDASVGLAMLAAGSMCLFIA 358
 324 vlsgvwmwfstsaafayfkltdgpgpnsahvasapvsagvdasvglawlagvmmclfia 383
 359 AGFAVAGWGPFPWLLMSEIFPLHVKGVANGICVLTNMLMAFLVTKFESSLMELVLRPGAFWL 418
 384 agfavgwgpfpwllmseifplhvkvgangicvltnwlmalfvltkesslmevlrpgafwl 443
 419 LASARCFISVLTFLFCVPERTKGKTLEOITAHFEGR 453
 444 lsaarcflsvltflfcvpeikgktleqitahfegr 478

RESULT 4
 AAB66933 standard; Protein: 478 AA.
 ID AAB66933 standard; Protein: 478 AA.
 AC AAB66933;
 DT 17-APR-2001 (first entry)
 DE Rat GLUTX1.
 KW Rat; GLUTX: gene therapy; vaccine; hexose transport modulator;
 KM hexose transport disorder; ischemia; diabetes; hyperglycaemia;
 hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 OS Rattus sp.
 PN WO200104145-A2.
 PD 18-JAN-2001.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 13:37:40 ; Search time 2842.4 Seconds
(without alignments)
1066.109 Million cell updates/sec

Title: US-09-516-493-11

Perfect score: 282

Sequence: 1 gagccgcgcgtatgttcaccc.....tccctcagcgccttccttcacg 282

Scoring table:

IDENTITY-NUC
Gapop 10.0 ; Gapext 1.0

Searched: 11351937 seqs, 537289281 residues

Total number of hits satisfying Chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estipl:*
6: em_estlba:*
7: em_estlro:*
8: em_estlov:*
9: em_hic:*
10: qb_estl:*
11: qb_estl2:*
12: qb_hic:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inu:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	279.4	99.1	319	11	H34451
2	225.6	80.0	823	11	BF120840
3	221.8	78.7	546	10	BE231636
4	219.2	77.7	949	11	BE231636
5	199.4	70.7	934	11	BE910478
6	188.2	66.7	712	11	BE290178
7	171.2	60.7	496	10	AM248655
8	168.8	59.9	440	11	BF774206
9	167.2	53.3	752	10	AL565390
10	154	54.6	538	11	BF146289
11	149.4	53.0	235	11	BF146289
12	149.4	53.0	327	11	BF742266

13	149.2	52.9	917	11	BF140667
14	137.2	48.7	595	11	BF079217
15	108.4	38.4	358	10	AA627408
16	102.8	36.5	295	11	BF303512
17	95.2	33.8	480	10	AM249090
18	89.4	31.7	359	10	AM141319
19	89.4	31.7	684	11	BF325106
20	89.4	31.7	912	11	BF206274
21	88.8	31.5	891	10	AL405787
22	77.4	27.4	911	11	BF311341
23	74	26.2	830	10	BF700749
24	67	23.8	974	10	AL534131
25	67	23.8	1079	13	CNS055KF
26	66.2	23.5	638	11	BF549214
27	62.2	22.1	289	11	C88387
28	61.8	21.9	279	11	H46443
29	61.2	21.6	564	11	BF388988
30	60.4	21.4	543	10	BE598634
31	60.4	21.4	555	10	AV524951
32	60.4	21.4	588	10	BE594919
33	60.4	21.4	595	10	BE596754
34	60.4	21.4	654	10	BE594918
35	59.4	21.1	525	10	AM463536
36	59.4	21.1	808	11	BF260051
37	59.2	21.0	855	10	AM940827
38	58.2	20.6	559	10	AM313298
39	56.4	20.0	459	10	BE482283
40	56.2	19.9	796	13	AA985752
41	56.2	19.9	803	10	BE130163
42	55.8	19.8	487	11	BF878576
43	55.4	19.6	795	13	AO324535
44	55.2	19.6	521	10	AV522458
45	55.2	19.6	562	10	AV526551

ALIGNMENTS

RESULT 1
H34451
LOCUS H34451 319 bp mRNA
DEFINITION EST11406 Rat PC-12 cells, NGF-treated (9 days) Rattus sp. cDNA 5' end similar to Glucose transporter, mRNA sequence.

ACCESSION H34451.1 GI:979868

VERSION H34451.1
KEYWORDS
SOURCE

ORGANISM

REFERENCE
Rattus sp.
Rattus sp.
Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

AUTHORS
1 (bases 1 to 319)
Lee,N.H., Welnsstock,K.G., Kirness,E.F., Earle-Hughes,J.A., Fuldner,R.A., Matmaras,S., Glodex,A., Gocayne,J.D., Adams,M.D., Kerlavage,A.R., Fraser,C.M. and Venter,J.C.

TITLE
Comparative expressed-sequence tag analysis of differential gene expression profiles in PC-12 cells before and after nerve growth factor treatment

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 92, 8303-8307 (1995)

MEDLINE
95396786

COMMENT
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@igf.org

For clone availability please contact the TIGR Database
(tdinfo@db.tigr.org)

Seg primer: M13 Reverse.

Location/Qualifiers

FEATURES
source

1..319
/organism="Rattus sp."
/db_xref="ATCC (inhost):2005510"

/db_xref="taxon:10118"
 /clone_lib="Rat PC-12 cells, NGF-treated (9 days)"
 /note="Vector: pBluescript SK⁺; Site_1: EcoRI; Site_2:
 XhoI; poly(A)⁺ RNA was purified from 9-day NGF treated
 PC12 cells. cDNA was constructed using an oligo-dT primer
 and directionally cloned using the Lambda Zap II Vector
 Kit by Stratagene"

BASE COUNT 48 a 102 c 80 g 88 t 1 others
 ORIGIN

Query Match 99.1%; Score 279.4; DB 11; Length 319;
 Best Local Similarity 99.3%; Pred No. 5.2e-61;
 Matches 280; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gagctcgtatgttaccctggcctggcctggcctgtgctgtagcagcatgtcccttc 60
 ||||||| ||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 DB 19 GAGCGTGTGATGTTACCGTGGGCTGGCTGGCTGTAGCAGCATGTCCCTTTC 78
 OY 61 atcgctgtttgcagtaaggctggggagccatccctgtgctcctcatgtcaagatcttc 120
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 DB 79 ATCGCTGTTTGCATAGCTGGGAGCCCATCCCTGGCTCCTCATGTGAGATCTTC 138
 OY 121 cctctgcacataaagggtgtggtaccgggtctgtgtcctcaacagtgtcatggcc 180
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 DB 139 CCTCTGCATCAAGGAGTGTGCTACCGGCTGTGTCTTCTCACCACTGGTTCAATGGC 198
 OY 181 ttctgtgtgaccaaagatttaacacatcatgtagatcctcagaccctcagcgccctc 240
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 DB 199 TTTTGTGTGACCAAGAGTTTACACATCATGTGAGATCCTCAGACCCTTACGGCGCTTC 258
 OY 241 tggctacacgcgtccttcgtatcctcctcagcgctccttcacg 282
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 DB 259 TGGCTCACCGCTGCTGTATCTCAGCGCTCTTTTCACG 300

RESULT 2
 LOCUS BF120840 823 bp mRNA EST 24-OCT-2000
 DEFINITION 601757442F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:398651 5',
 mRNA sequence.
 ACCESSION BF120840
 VERSION BF120840.1 GI:10959880
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 823)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL, at:
 http://image.llnl.gov
 Plate: LAM9192 row: d column: 12
 High quality sequence start: 4
 High quality sequence stop: 644.
 Location/Qualifiers
 1..823
 /organism="Mus musculus"
 /taxon="10090"
 /db_xref="taxon:10090"
 /clone_lib="CGAP_Mam5"
 /tissue="gross tissue"

/lab_host="DH10B"
 /note="Organ: mammary; Vector: PCMV-SPOB6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Robin Humphreys,
 NIH"
 BASE COUNT 150 a 226 c 257 g 190 t
 ORIGIN

Query Match 80.0%; Score 225.6; DB 11; Length 823;
 Best Local Similarity 90.9%; Pred. No. 2.8e-47;
 Matches 240; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

OY 1 gagctcgtatgttaccctggcctggcctggcctgtgctgtagcagcatgtcccttc 60
 ||||||| ||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 DB 204 GAGCGTGTGATGTTACCGTGGGCTGGCTGGCTGTAGCAGCATGTCCCTTTC 263
 OY 61 atcgctgtttgcagtaaggctggggagccatccctgtgctcctcatgtcaagatcttc 120
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 DB 264 ATTCGCTGTGGGAGTGTGGCTGGGAGCCCATCCCTGGCTCCTCATGTGACAGATCTTC 323
 OY 121 cctctgcacataaagggtgtggtaccgggtctgtgtcctcaacagtgtcatggcc 180
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 DB 324 CCTCTGCATGTCAAGGAGTGTGCTACCGGCTGTGTCTTCTCACCACTGGTTCAATGGC 383
 OY 181 ttctgtgtgaccaaagatttaacacatcatgtagatcctcagaccctcagcgccctc 240
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 DB 384 TTTTGTGTGACCAAGAGTTTACACATCATGTGAGATCCTCAGACCCTTACGGCGCTTC 443
 OY 241 tggctacacgcgtccttcgtatcctcctcagcgctccttcacg 264
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 DB 444 TGGCTCACCGCTGCTGTATCTCAGCGCTCTTTTCACG 467

RESULT 3
 LOCUS BE231636 546 bp mRNA EST 10-JUL-2000
 DEFINITION 136437 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION BE231636
 VERSION BE231636.1 GI:9016354
 KEYWORDS EST.
 SOURCE pig.
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 546)
 Fahnenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
 Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.
 and Keele, J.W.
 Design and use of two pooled tissue normalized cDNA libraries for
 EST discovery in swine
 Unpublished (2000)
 JOURNAL Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt-trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.
 PCR PRIMERS
 FORWARD: AGGAACACGTATGACCAT
 BACKWARD: GTTTCGAGTCACGAGC
 Plate: 70 row: A column: 21
 Seq primer: ATTAGGTGACACTATAG.
 Location/Qualifiers
 1..546
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="MARC 1P1G"
 /tissue_type="pooled"

FEATURES
 Source

BASE COUNT 144 07:44:19 2002
 /lab,
 /tissue="gross tissue"

DB	303	TGGAACAAATCACAGCCCATTTTGAGGGGGGAGACAGCCAC	343
RESULT	13		
LOCUS	H34372	324 bp	mRNA
DEFINITION	EST11272 Rat PC-12 cells, NGF-treated (9 days) Rattus sp. CDNA clone RPNBT28 3' end, mRNA sequence.		
ACCESSION	H34372		
VERSION	H34372.1	GI:979789	
KEYWORDS	EST.		
SOURCE	Rattus sp.		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 324)		
AUTHORS	Lee, N.H., Weinstein, K.G., Kirkness, E.F., Earle-Hughes, J.A., Fuldner, R.A., Mammas, S., Glodek, A., Gocayne, J.D., Adams, M.D., Kerlavage, A.R., Fraser, C.M. and Venter, J.C.		
TITLE	Comparative expressed-sequence-tag analysis of differential gene expression profiles in PC-12 cells before and after nerve growth factor treatment		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 92, 8303-8307 (1995)		
MEDLINE	9536786		
COMMENT	Contact: Lee, NH The Institute for Genomic Research 9712, Medical Center Drive, Rockville, MD 20850, USA Tel: (301)-838-3529 Fax: (301)-838-0208 Email: nhlee@tigr.org For clone availability please contact the TIGR Database (tldinfo@tigr.tigr.org) Seq primer: M13 - 21.		
FEATURES	Location/Qualifiers		
source	1..324		
	/organism="Rattus sp."		
	/db_xref="ATCC (Inhost):2005377"		
	/db_xref="taxon:10118"		
	/clone="RPNBT28"		
	/clone_11b="Rat PC-12 cells, NGF-treated (9 days)"		
	/note="Vector: Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; poly(A)+ RNA was purified from 9-day NGF treated PC12 cells. cDNA was constructed using an oligo-dT primer and directionally cloned using the Lambda Zap II Vector Kit by Stratagene"		
BASE COUNT	89 a 74 c 79 g 78 t 4 others		
ORIGIN			
Query Match	22.1%	Score 229	DB 11; Length 324;
Best Local Similarity	95.5%	Pred. No. 4, 7e-22;	
Matches 298; Conservative	0;	Mismatches 8;	Indels 6; Gaps 6;
QY 670	tttgggtgcttggtgcttaagaac-tggccatactc-acttgatctgggatgaagaa- 726		
DB 314	tttgggtgcttggtgcttaagaac-tggccatactc-acttgatctgggatgaagaa- 255		
QY 727	aggagacttgcacataaagaatttgggtcagaagaagaagtcagtgatccagaagaa 786		
DB 254	aggagacttgcacataaagaatttgggtcagaagaagaagtcagtgatccagaagaa 195		
QY 787	agagaatggtctctgtctgttcaaccagaagtccttctcaagatg-ccaaagacctccgat 845		
DB 194	agagaat-gttcttgccttggtcacaacagtccttctcaagatg-ccaaagacctccgat 136		
QY 846	tcaacttgggttagtgcacattaccataacttcaaggttcttccaaacttcaagctgat 905		

Db	135	TCACCTGGGGGTAAAGCCAGTTACCCACNNACTTAAAGATTCTCTCCCAACTCTTCCCGGT	76
Qy	906	ctcagtcctcggatcatatagtcaccaggtctgtgtgatttcagaaataaaggct	965
Db	75	CTCACTGTCTCGGATCATATTAGTACACAGGCTTGTGTGATTTCAG - AAAATTAAGGCT	17
Qy	966	cttccgttcaa 977	
Db	16	CTTCCGTCAA 5	
RESULT	14		
LOCUS	AL565390/c		
DEFINITION	AL565390 LTI_FLO13_FBrn1 Homo sapiens cDNA clone CS0DF05YE02 3	EST	16-FEB-2001
ACCESSION	AL565390	prime, mRNA sequence.	
VERSION	AL565390.1	GI:12916718	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
FEATURES			
Source	Location/Qualifiers		
	1..752		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0DF005YE02"		
	/clone_lib="LTI_FLO13_FBrn1"		
	/dex_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"		
	/lab_host="DH10B"		
	/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifestech.com URL : http://fulllength.invitrogen.com"		
BASE COUNT	177 a 206 c 242 g 118 t	9 others	
ORIGIN			
Query Match	21.1%, Score 218.6; DB 10; Length 752;		
Best Local Similarity	79.5%; Pred. No. 8.4e-21;		
Matches	290; Conservative 2; Mismatches 67; Indels 6; Gaps 3;		
Qy	57	gtgtccatctccgaagagcctgcgtgatgtcactctgtggcctgtgcctgtgcgtgtagg	116
Db	731	SGGGCTGTCTCTGTGCACAGCGCTGTGAGCCAGCGCTGGGCTSGCTTGCGTGGCGTGN	672
Qy	117	cagatgtgcctctcatc---cgtcgttttgcagtaagctcggagaccatcccc-tggc	172
Db	671	CANCAATGTGCTCTTTCANATGCCCNAGCTTTTGCGGTGGGCTGGGGGCCCATCCCTTGGC	612
Qy	173	tctcatatgcagagatctt--cctctgcacatcaagaagggttgcatacggcgtctgtgt	230
Db	611	TTCCTATGTGACAGATWTTCCGCCCHGCACATNTCAAGGGCGTGGGACACAGGATCTGCGT	552
Qy	231	cctaaccaactggttcataagccttctcgtgtgacccaagaagtttaacagcatcatgagat	290
Db	551	CCNACCAAGATGGCTCAATGGGCTTTCGTGTATACCAAGAGTTCAAGCAACCATATGAGAGT	492

Query Match 14.5%: Score 41: DB 6: Length 2217;
 Best Local Similarity 100.0%: Pred. No. 9, 1e-12;
 Matches 41: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Y 88 cccatccctgctccatgctcagagatcttcctctgca 128
 1518 CCCATCCCTGCTCCTCATCTCAGACATCTCCCTCTGCA 1558

RESULT 13

AL445222

LOCUS 225370 bp DNA PRI 24-APR-2001
 DEFINITION Human DNA sequence from clone RP11-356B19 on chromosome 9, complete
 sequence.

ACCESSION

AL445222

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 225370)

Laird, G.

Direct Submission

Submitted (24-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

requests: clonequests@sanger.ac.uk

On Apr 26, 2001 this sequence version replaced gi:13277497.

During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations.variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: EM, EMBL; SW,
 SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpepThis sequence was generated from part of bacterial clone contigs of human
 chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/MGP/Chr9RP11-356B19 is from the library RPCI-11.2 constructed by the group
 of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm

VECTOR: pBAC3.6

This sequence is the entire insert of clone RP11-356B19. The true
 left end of clone RP11-373J8 is at 96439 in this sequence. The true
 right end of clone RP13-225021 is at 9980 in this sequence.

FEATURES

source

1. 225370

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="9"

/clone="RP11-356B19"

/clone_lib="RPCI-11.2"

23455..23821

misc_feature

restriction digest data.

/note="Single clone region. Assembly confirmed by
 restriction digest data."

misc_feature

restriction digest data confirm the assembly."

misc_feature

225123..225370

/note="Single clone region. Sequence from reads from a
 short insert library derived from a single pUC clone.
 Restriction digest data confirm the assembly."

BASE COUNT 56329 a 57092 c 57267 g 54682 t
 ORIGIN

Y 88 cccatccctgctccatgctcagagatcttcctctgca 128
 Db 61234 CCCATCCCTGCTCCTCATCTCAGACATCTCCCTCTGCA 61274

RESULT 14

AX191507

LOCUS AX191507 1461 bp DNA PAT 15-AUG-2001
 DEFINITION Sequence 29 from Patent WO0149728.
 ACCESSION AX191507
 VERSION AX191507.1 GI:15209697
 KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

ORIGIN

195 a

501 c

456 g

309 t

Query Match

Best Local Similarity

Matches

23: Conservative

0: Mismatches

0: Indels

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

0: Gaps

AUTHORS Gale, J.M., Tobey, R.A. and D'Anna, J.A.
 TITLE Localization and DNA sequence of a replication origin in the
 JOURNAL rhodopsin gene locus of Chinese hamster cells
 MEDLINE J. Mol. Biol. 224 (2), 343-358 (1992)
 REFERENCE 92219256
 2 (bases 1 to 11931)
 AUTHORS D'Anna, A.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUL-1991) A. D'Anna, Los Alamos National Lab, PO Box
 1663/MS-M880, Los Alamos, NM 87545, USA

FEATURES
 source Location/Qualifiers
 1. 11931
 /organism="Cricetulus griseus"
 /db_xref="taxon:10029"
 /tissue_type="ovary"
 /cell_line="CHO"
 /clone_id="cosmid: scos1"
 3269..3426
 /join(6107..6467,7929..8097,9177..9342,9455..9694,
 10702..10812)
 /gene="rhodopsin"
 6107..10812
 /gene="rhodopsin"
 /join(6107..6467,7929..8097,9177..9342,9455..9694,
 10702..10812)
 /gene="rhodopsin"
 /codon_start=1
 /product="opsin"
 /protein_id="CAA43398.1"
 /db_xref="GI:49479"
 /db_xref="SWISS-PROT:P28681"
 /translation="MNGTEGPNFYVPESNATGVVRSPEFYPOVYLAEPWQFSLAAYM
 FLIVLGEPINFLVTVQHKLRLPLNYILLNLAVALFMVFGFTTLITSLHG
 FVEGPGCNLEGFATLGGELALMSLVLAIERVVICPKPMNFRGEMHAIMCVET
 WMLAACAPLVGMSRYPGSGCQGVDTLKPEVNNSEFVIMFVETIPLIV
 IFFCYGLVFTVKEAQAQODESATTOKAREKVTBVLIMVFFLICMPVAGVAFLYF
 THQSNFGPIFMTLPAPFAKSSITNPVITIMNKRFRNCMLTTLCCKNLTGDDKAS
 ATASKTETSQVAPA"
 prim_transcript <6107..>10812
 /gene="rhodopsin"
 join(<6107..6467,7929..8097,9177..9342,9455..9694,
 10702..>10812)
 /gene="rhodopsin"
 <6107..6467
 /gene="rhodopsin"
 /number=1
 6468..7928
 /gene="rhodopsin"
 /number=1
 6944..7070
 /rpt_family="Alu"
 7929..8097
 /gene="rhodopsin"
 /number=2
 8098..9176
 /gene="rhodopsin"
 /number=2
 9177..9342
 /gene="rhodopsin"
 /number=3
 9343..9454
 /gene="rhodopsin"
 /number=3
 9455..9694
 /gene="rhodopsin"
 /number=4
 9695..10701
 /gene="rhodopsin"
 /number=4
 10031..10159
 /rpt_family="Alu"
 10702..>10812
 /gene="rhodopsin"

repeat_region /number=5
 11409..11462
 /note="triplex motif (GA)"
 BASE COUNT 2719 a 3172 c 3122 g 2918 t
 ORIGIN

Query Match 7.4%; Score 21; DB 10; Length 11931;
 Best Local Similarity 100.0%; Pred.No. 1.2; 0; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches

Qy	32	ggctgctgtaggaacatgt	52
Db	591	GGCTGCTGTAGGCAAGCATGT	611

Search completed: February 13, 2002, 20:07:08
 Job time: 23045 sec

07:44:06

SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 2085)
TITLE	Pfeundschuh,M.
JOURNAL	Hodgkin's disease associated molecules and uses thereof
FEATURES	Patent: US 5840568-A 7 24-NOV-1998;
source	location/Qualifiers
	1..2085
BASE COUNT	/organism="unknown"
ORIGIN	746 a 398 c 557 g 384 t
Query Match	5.5%; Score 57; DB 6; Length 2085;
Best Local Similarity	100.0%; Pred. No. 2e-21;
Matches	57; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	1 tggcgcgcctctagaactagtgatcccccggcgctcaggaaattcgacagcgt 57
Dn	16 TGGCGCCGCTCTAGACTAGTGATGCCCGCGGCTGCAGGAATTGCGACGCTG 72
RESULT	9
LOCUS	AX062918 380 bp DNA PAT 24-JAN-2001
DEFINITION	Sequence 545 from Patent WO0100828.
ACCESSION	AX062918
VERSION	AX062918.1 GI:12540776
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 380) Wang,T., Banur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Retter,M.W. and Mannion,J. Compositions and methods for the therapy and diagnosis of Lung Cancer Patent: WO 0100828-A 545 04-JAN-2001; CORIXA CORPORATION (US)
TITLE	Location/Qualifiers
JOURNAL	1..380
FEATURES	/organism="Homo sapiens" source /db_xref="taxon:9606" 1..380
misc_feature	/note="n = A,T,C or G"
BASE COUNT	100 a 76 c 102 g 97 t 5 others
ORIGIN	
Query Match	5.4%; Score 56; DB 6; Length 380;
Best Local Similarity	100.0%; Pred. No. 8.4e-21;
Matches	56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 tggcgcgcctctagaactagtgaatccccggcgctcaggaaattcgacagcgt 56
Dn	332 TGGCGCCGCTCTAGACTAGTGAATCCCCCGGCTGCAGGAATTGCGACGAGCT 277
RESULT	10
LOCUS	AB005878 518 bp mRNA PLN 25-JUL-1997
DEFINITION	Nicotiana tabacum mRNA for BYU15, partial cds.
ACCESSION	AB005878
VERSION	AB005878.1 GI:2280517
KEYWORDS	BYU15.
SOURCE	Nicotiana tabacum (strain:bright yellow 2) cDNA to mRNA.
ORGANISM	Nicotiana tabacum
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asterales; euasterids I; Solanales; Solanaceae; Nicotiana. 1 (bases 1 to 518)
REFERENCE	

AUTHORS Imanishi, S.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-1997) to the DDBJ/EMBL/GenBank databases.
Shunsuke Imanishi, School of Agriculture, Nagoya University,
Laboratory of Biochemistry, Furo-cho, Chikusa, Nagoya 464-01, Japan
(E-mail: i45184@nuc.cc.nagoya-u.ac.jp, Tel: 052-789-4098,
Fax: 052-789-4094)

REFERENCE
AUTHORS 2 (sites)
TITLE Kojima, H., Hashizume, K., Imanishi, S. and Nakamura, K.
JOURNAL Early Jasmonate-Inducible genes of Tobacco Cells
FEATURES Unpublished (1997)
location/Qualifiers
1..518
/organism="Nicotiana tabacum"
/strain="Bright yellow 2"
/db_xref="taxon:4097"
<1..>518
/codon_start=2
/product="BYJ15"
/protein_id="BA21615.1"
/db_xref="GI:2280518"
/translation="LELHGGGSRSTSGSPGLQEFGTSTFFNRPTGRFNGRIIDF
ITOSFKLPFLNAYLDGAFTGAVAGGATLNTSPMTARNIRLPTWNTPLANLG
WFKTHOSTCGSRCADSLKSLIVMGEMGNDYNGFPQKJISEVRTYYPNVVAGIM
RGTRCDSARSH"

BASE COUNT 121 a 142 c 122 g 133 t
ORIGIN

Query Match 5.4%; Score 56; DB 8; Length 518;
Best Local Similarity 100.0%; Pred. No. 8.3e-21;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGGCGGCGCTCTAGACTAGTGGATGATCCCCGGGCTGCAAGAAATCGCAGCAGCT 56
|||||
Db 19 TGGCGGCGCTCTAGACTAGTGGATGATCCCCGGGCTGCAAGAAATCGCAGCAGCT 74
|||||

RESULT 11
AB002449 1228 bp mRNA PRT 13-FEB-1999
LOCUS Homo sapiens mRNA from chromosome 5q21-22, clone:843EX.
DEFINITION AB002449
VERSION AB002449.1 GI:2943812
KEYWORDS advanced lung cancer.
SOURCE Homo sapiens tissue_lib:fetal brain cDNA to mRNA, clone:843EX.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1228)
Ueno, K.
Direct Submission
Submitted (29-MAR-1997) to the DDBJ/EMBL/GenBank databases.
Kiyonobu Ueno, Osaka University Medical School, Department of
Medicine III, 2-2, Yamada-oka, Suita, Osaka, Japan, Suita, Osaka
565, Japan (E-mail: ueno-k@kinchu.hosp.go.jp, Tel: 06-879-3835,
Fax: 06-879-3839)
2 (sites)
Ueno, K., Kumagai, T., Kijima, T., Kishimoto, T. and Hosoe, S.
Cloning and tissue expression of cDNAs from chromosome 5q21-22
which is frequently deleted in advanced lung cancer
Hum. Genet. 102 (1), 63-68 (1998)
98141129
location/Qualifiers
1..1228
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="843EX"
/map="5q21-22"
/tissue_lib="fetal brain"
<1..>1228
/note="expressed ubiquitously"

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
FEATURES
source

BASE COUNT 415 a 225 c 205 g 377 t 6 others
ORIGIN

Query Match 5.4%; Score 56; DB 9; Length 1228;
Best Local Similarity 100.0%; Pred. No. 7.9e-21;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGGCGGCGCTCTAGACTAGTGGATGATCCCCGGGCTGCAAGAAATCGCAGCAGCT 56
|||||
Db 31 TGGCGGCGCTCTAGACTAGTGGATGATCCCCGGGCTGCAAGAAATCGCAGCAGCT 86
|||||

RESULT 12
AX101122 1388 bp DNA PAT 10-APR-2001
LOCUS Sequence 1 from Patent WO0121650.
DEFINITION AX101122
ACCESSION AX101122
VERSION AX101122.1 GI:13619973
KEYWORDS
SOURCE Dimorphotheca.
ORGANISM Dimorphotheca
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; eusterids II; Asterales; Asteraceae; Asteroideae;
Calenduleae.
1 (bases 1 to 1388)
Coldren, C., Flint, D., Hallahan, D.L. and Wang, H.
Cis-prenyltransferases from plants
Patent: WO 0121650-A.1 29-MAR-2001;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
location/Qualifiers
1..1388
/organism="Dimorphotheca"
/db_xref="taxon:13490"

BASE COUNT 437 a 268 c 297 g 386 t
ORIGIN

Query Match 5.4%; Score 56; DB 6; Length 1388;
Best Local Similarity 100.0%; Pred. No. 7.8e-21;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGGCGGCGCTCTAGACTAGTGGATGATCCCCGGGCTGCAAGAAATCGCAGCAGCT 56
|||||
Db 231 TGGCGGCGCTCTAGACTAGTGGATGATCCCCGGGCTGCAAGAAATCGCAGCAGCT 286
|||||

RESULT 13
I14842 114842 1601 bp DNA PAT 02-APR-1996
LOCUS Sequence 10 from patent US 5455167.
DEFINITION I14842
ACCESSION I14842
VERSION I14842.1 GI:1249751
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
1 (bases 1 to 1601)
Voelker, T.A. and Davies, H.M.
Medium-chain thioesterases in plants
Patent: US 5455167-A 10 03-OCT-1995;
location/Qualifiers
1..1601
/organism="unknown"

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
FEATURES
source

BASE COUNT 375 a 372 c 423 g 425 t
ORIGIN

Query Match 5.4%; Score 56; DB 6; Length 1601;
Best Local Similarity 100.0%; Pred. No. 7.7e-21;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2002, 13:37:40 ; Search time 2842.4 Seconds

(without alignments)
1066.109 Million cell updates/sec

Title: US-09-516-493-11

Sequence: 1 gagcctgctgctgcttccaccc.....tccctcagcgtcccttccacg 282

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estda:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_estl:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	279.4	99.1	319	11	H34451
2	225.6	80.0	823	11	BF120840
3	221.8	78.7	546	11	BE231636
4	219.2	77.7	949	11	BE749509
5	199.4	70.7	934	11	BE910478
6	188.2	66.7	712	11	BE290178
7	171.2	60.7	496	10	AM248655
8	168.8	59.9	440	11	BF774206
9	167.2	59.3	752	10	AL565390
10	154	54.8	538	11	BF7742321
11	149.4	53.0	235	11	BF146289
12	149.4	53.0	327	11	BF742266

13	149.2	52.9	917	11	BF140667
14	137.2	48.7	595	11	BF079217
15	108.4	38.4	358	11	AA627408
16	102.8	36.5	295	11	BE303512
17	95.2	33.8	480	10	AM249090
18	89.4	31.7	359	10	AM141319
19	89.4	31.7	684	11	BE325106
20	89.4	31.7	912	11	BE206274
21	88.8	31.5	891	11	AI405787
22	77.4	27.4	911	11	BF113441
23	74	26.2	830	11	BE700749
24	67	23.8	974	10	AL534131
25	67	23.8	1079	13	CNS055KF
26	66.2	23.5	638	11	BE549214
27	62.2	22.1	289	11	C88387
28	61.8	21.9	279	11	H46443
29	61	21.6	564	11	BE388988
30	60.4	21.4	543	10	BE598634
31	60.4	21.4	555	10	AV524951
32	60.4	21.4	588	10	BE594919
33	60.4	21.4	595	10	BE596754
34	60.4	21.4	654	10	BE594918
35	59.4	21.1	525	10	AM463536
36	59.4	21.1	808	11	BE260051
37	59.2	21.0	855	10	AM940827
38	58.2	20.6	560	10	AM913298
39	56.4	20.0	459	10	BE482283
40	56.2	19.9	796	13	AZ985752
41	56.2	19.9	803	11	BE130163
42	55.8	19.8	487	11	BE878576
43	55.4	19.6	795	13	AQ324535
44	55.2	19.6	521	10	AV522458
45	55.2	19.6	562	10	AV526551

ALIGNMENTS

RESULT 1
H34451
LOCUS H34451 319 bp mRNA
DEFINITION EST111406 Rat PC-12 cells, NGF-treated (9 days) Rattus sp. CDNA 5' end similar to Glucose transporter, mRNA sequence.

ACCESSION H34451
VERSION H34451.1
KEYWORDS GI:979868
SOURCE EST.
ORGANISM Rattus sp.
Rattus sp.
Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 319)
Lee N.H., Weinstein K.G., Kirkness E.F., Earle-Hughes J.A., Fulmer R.A., Marmaras S., Giodok A., Gocayne J.D., Adams M.D., Kerlavage A.R., Fraser C.M. and Venter J.C.
Comparative expressed-sequence-tag analysis of differential gene expression profiles in PC-12 cells before and after nerve growth factor treatment
Proc. Natl. Acad. Sci. U.S.A. 92, 8303-8307 (1995)

JOURNAL MEDLINE

COMMENT 95396786
Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@igf.org
For clone availability please contact the TIGR Database (tdbinfo@tdb.tigr.org)
Seq primer: M3 Reverse.

FEATURES
source
1..319
/organism="Rattus sp."
/db_xref="ATCC (inhost):2005510"

P70 Seq Search Report


```
/organism="Sus scrofa"  
/db_xref="taxon:9823"  
/clone_1fb="MARC 1P1G"  
/tissue_type="pooled"
```


Accession	Sequence	Position
QY	cttggggggtcgtgtcccaagcaac- <u>ggccataacc</u> -acttgaccggggatgagaa-	760
QY	670	
Db	tttggggggtttggacccttaagcaacttgcocatcatccnacttctactcgggggatgaaaa	255
Db	314	
QY	aggacattagccacataagatcttagcgtccaaacaaggtccagtlgtgtccagaaagaa	766
QY	727	
Db	254	
QY	agagaatgtctctgtctctgcaaccaagttccttcctagatg- <u>cgaagacc</u> ctccgat	845
QY	787	
Db	134	
QY	846	
QY	taacttggttttagccagcttaccacatcacttcaagttctctccaactctcagctggt	905

QY	57	ggttccatctccgcagagccgtctgtatgtttcacccttggtgctccgtctgctgaagg	116
Db	731	SGGCGCTGTCTCTCCACAGCGCTGTGATGCGCAGCGTGGGGCTSGCCCTGGGTGGCGCGGN	672
QY	117	cagcatgtgectctcat--cgtgtatttgcagtaagctctgtagaccatccccc-tygc	172
Db	671	CANATGTGCGCTTTTGNATNGCGCGGCTTTTGGGATGGGCTGGGGGCGCCATCCCTTGGC	612
QY	173	tctcatgttcaagatcctt--ccctctgcacatcaagaagtgtgtctacccgcgtctgtgt	230
Db	611	TCTCATGTCTAGAGATCTTCCCCCTCCGACATMTCGAAGGGCTGGCGACAGGCATMTCGT	552
QY	231	ccttcaacaaactgttcaatgagccttctctgtgtgacccaagaagtttaacagcatcatggaac	290
Db	551	CCNACCACAAAGCGCTCATGTGGCTTTTCTCGGACCAAGAGATTTTAGCGAGCTCTATGGAGGT	492

AUTHORS Gale, J.M., Tobey, R.A. and D'Anna, J.A.
 TITLE Localization and DNA sequence of a replication origin in the rhodopsin gene locus of Chinese hamster cells
 JOURNAL J. Mol. Biol. 224 (2), 343-358 (1992)
 MEDLINE 92219256
 REFERENCE 2 (bases 1 to 11931)
 AUTHORS D'Anna, A.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUL-1991) A. D'Anna, Los Alamos National Lab, PO Box 1663/MS-M880, Los Alamos, NM 87545, USA

FEATURES
 Source Location/Qualifiers
 1..11931
 /organism="Cricetulus griseus"
 /db_xref="taxon:10029"
 /tissue="type="Ovary"
 /cell_line="CHO"
 /clone_lib="cosmid: scos1"
 repeat_region 3269..3426
 join(6107..6467,7929..8097,9177..9342,9455..9694,10702..10812)
 /gene="rhodopsin"
 6107..10812
 /gene="rhodopsin"
 join(6107..6467,7929..8097,9177..9342,9455..9694,10702..10812)
 /gene="rhodopsin"
 /codon_start=1
 /product="opsin"
 /protein_id="CAA43398.1"
 /db_xref="GI:49479"
 /translation="MNGTEGPNFYVPFSNATGVRSPEFYPOYLAEPWQFSMLAAYMFLIIVGEPINFLTYVQHKLRPLNYILNLAVADLEWFGGFTTLTSLHGYPVFGPTGNIEGFATLGEIALLMSLVLAIEDRYVVICPKMSNFRGENHAIMGVFTWIMALACAPPLVKSAAQYIPESMOCSCGVDTYTLKPEVNNESEFVIYMFVYHFTPLIIFPCYGLVFTVKEAAQOQESATTOKAKEVTRMYILWVFEFLICMPYAGVAFYIVTHGSGNFGPIEMTLPAPFAKSSIVNPVIYIMNKNQFRNCMLTTLCCGKNITGDDEASATASKTETSQVAPA"
 prim_transcript <6107..>10812
 /gene="rhodopsin"
 join(<6107..6467,7929..8097,9177..9342,9455..9694,10702..>10812)
 /gene="rhodopsin"
 <6107..6467
 /gene="rhodopsin"
 /number=1
 6468..7928
 /gene="rhodopsin"
 /number=1
 7929..8097
 /gene="rhodopsin"
 /number=2
 8098..9176
 /gene="rhodopsin"
 /number=2
 9177..9342
 /gene="rhodopsin"
 /number=3
 9343..9454
 /gene="rhodopsin"
 /number=3
 9455..9694
 /gene="rhodopsin"
 /number=4
 9695..10701
 /gene="rhodopsin"
 /number=4
 10031..10159
 /rpt_family="Alu"
 10702..>10812
 /gene="rhodopsin"

/number=5
 repeat_region 11409..11462
 /note="triplex motif (GA)"
 BASE COUNT 2719 a 3172 c 3122 g 2918 t
 ORIGIN

Query Match 7.4%; Score 21; PR 10; Length 11931;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 32 GCCTGCTGTAGCGCATGT 52
 |||||
 Db 591 GCCTGCTGTAGCGCATGT 611

Search completed: February 13, 2002, 20:07:08
 Job time: 23045 sec

07:44:06

/product="myosin heavy chain"
/protein_id="AB29713.2"
/db_xref="GI:8250661"
/translation="PPCCNSAGKLEGDLKLAOESTMDYENDKQQLDEKKKEFEN
SNLSQKLEDEQALQMLQKKIKELQARIIELEIEKASRAKAEKQSDLSLEEE
ISERLEBAGATSAOIEEMKREAEFOKMRLEAEATLOHEATATLRKHADSVLE
GEOIDNLOVYKQLEKSEKMEIDDLASMEVYSKSKNLEKCTLEDQVSELT
KEEEOORLIELTQORGLQTEGSEYRQIDEDLSVLSQSRQKATQOIEELKROL
EEVKAASALAHQSRHCDLRLREYEEDEKAELOKAMSANSEVAQWRTKET
DAIORTLELEAKKLAORLQDAEHEVAENACASLEKTKORLQNEVEDLIDVERT
NAACALDKQKRNFDKILAEWKQYETHEAELEASQESKSLTELEKIKNAEESD
QLETKRNNKNOEISDLTEQIAEGKRTHIEKIKQYOEKSELOALFEAEASL
EHEGKILRIQLENOVSEIDRIKAEKDEIDQKRNHTRVSMSTDALEBSN
DAIRIKKMEGDLNEMEIOLNHSNMAALELRYRNQGLKOTLHLDLALQOEDL
KEOLAVERRANLQAEIEELRATLEQTERSKIADELIDASEVQLHTQNTSLIN
TKKLETDISOIGEBEDIVQENAEKKAITDAAMAEELKKEODTSALHEBK
KNEQVTKDLQRLDEADQALGKQKQJQKLEAVRELEGEVENEQKRVPAIKGR
KHRRVVELTYQTEEDRKNYRLQDLVDKLOSQKAVKQAEAEQSNVNLAKPKFI
QHELEAEERADIAESQVNLKRVKSRVHRTIIEE"

Query Match 6.0%; Score 62; DB 10; Length 2924;
Best Local Similarity 100.0%; Pred. No. 2.7e-24;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 976 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagggggg 1035
Db 2859 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgagggggg 1035

Qy 1036 cc 1037
Db 2919 CC 2920

RESULT 7
LOCUS AR070327 2605 bp DNA PAT 18-FEB-2000
DEFINITION Sequence 4 from patent US 5892010.
ACCESSION AR070327
VERSION AR070327.1 GI:7221215
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2605)
Gray, J., Collins, C., Hwang, S., Godfrey, T., Kowbel, D. and Kommens, J.
Genes from the 20013 amplicon and their uses
Patent: US 5892010-A 4 06-APR-1999;
FEATURES
Location/Qualifiers
1..2605
/organism="unknown"

BASE COUNT 698 a 625 c 550 g 732 t
ORIGIN

Query Match 5.6%; Score 58; DB 6; Length 2605;
Best Local Similarity 100.0%; Pred. No. 5.4e-22;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tggcgagcgctctagaactagatggtatcccccgggctgcgaagaattcgacagactg 58
Db 55 tggcgagcgctctagaactagatggtatcccccgggctgcgaagaattcgacagactg 112

RESULT 8
LOCUS AR060385 2085 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 7 from patent US 5840568.
ACCESSION AR060385
VERSION AR060385.1 GI:5986835
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2085)
Pfeundschuh, M.
Hodgkin's disease associated molecules and uses thereof
Patent: US 5840568-A 7 24-NOV-1998;
FEATURES
Location/Qualifiers
1..2085
/organism="unknown"
BASE COUNT 746 a 398 c 557 g 384 t
ORIGIN

Query Match 5.5%; Score 57; DB 6; Length 2085;
Best Local Similarity 100.0%; Pred. No. 2e-21;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tggcgagcgctctagaactagatggtatcccccgggctgcgaagaattcgacagactg 57
Db 16 tggcgagcgctctagaactagatggtatcccccgggctgcgaagaattcgacagactg 72

RESULT 9
LOCUS AX062918/c 380 bp DNA PAT 24-JAN-2001
DEFINITION Sequence 545 from Patent WO0100828.
ACCESSION AX062918
VERSION AX062918.1 GI:12540776
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 380)
Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,
Carter, D., Retter, M.W. and Mannion, J.
Compositions and methods for the therapy and diagnosis of lung
cancer
Patent: WO 0100828-A 545 04-JAN-2001;
FEATURES
Location/Qualifiers
1..380
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc-feature 1..380
/note="n = A,T,C or G"
BASE COUNT 100 a 76 c 102 g 97 t 5 others
ORIGIN

Query Match 5.4%; Score 56; DB 6; Length 380;
Best Local Similarity 100.0%; Pred. No. 8.4e-21;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tggcgagcgctctagaactagatggtatcccccgggctgcgaagaattcgacagactg 56
Db 332 tggcgagcgctctagaactagatggtatcccccgggctgcgaagaattcgacagactg 277

RESULT 10
LOCUS AB005878 518 bp mRNA PLN 25-JUL-1997
DEFINITION Nicotiana tabacum mRNA for BYJ15, partial cds.
ACCESSION AB005878
VERSION AB005878.1 GI:2280517
KEYWORDS BYJ15.
SOURCE Nicotiana tabacum (strain:bright yellow 2) cDNA to mRNA.
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; easterids I; Solanales; Solanaceae; Nicotiana.
REFERENCE 1 (bases 1 to 518)

AUTHORS Imanishi,S.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-1997) to the DDBJ/EMBL/GenBank databases.
Shunshuke Imanishi, School of Agriculture, Nagoya University,
Laboratory of Biochemistry, Furo-cho, Chikusa, Nagoya 464-01, Japan
(E-mail: i45184a@nuc.cc.nagoya-u.ac.jp, Tel: 052-789-4098,
Fax: 052-789-4094)
2 (sites)
REFERENCE Kojima,H., Hashizume,K., Imanishi,S. and Nakamura,K.
TITLE Early Jasmonate-Inducible genes of Tobacco Cells
JOURNAL Unpublished (1997)
FEATURES
source
location/Qualifiers
1..518
/organism="Nicotiana tabacum"
/strain="Bright yellow 2"
/db_xref="taxon:4097"
<1..>518
/codon_start=2
/product="BYJ15"
/protein_id="BA21615.1"
/db_xref="GI:2280518"
/translation="LELRGGGSRRTSGSPGLQEGTSSTFENRPTGRFSNGRYILDF
ITQSEKLEPLNAYLDKGAAPFGVNFAYGGTALTNSPTNRIPLPNNPLANLQ
MFKRLQSTGCSRCADSLKNSLIWGMGNDYNGFPQNKQISEVRYVNVAGIM
RGKRCDSARSH"
BASE COUNT 121 a 142 c 122 g 133 t
ORIGIN
Query Match 5.4%; Score 56; DB 8; Length 518;
Best Local Similarity 100.0%; Pred. No. 8.3e-21;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tggcgccgctctagaactagtgatcccccggcgctgcagaatcggcagcagct 56
|||||
DB 19 TGGCGCGCGCTCTAGAACTAGTGATCCCGCGGCTGCAGAAATTCGCGACGACT 74
RESULT 11
AB002449 1228 bp mRNA PRI 13-FEB-1999
LOCUS Homo sapiens mRNA from chromosome 5q21-22, clone:843Ex.
DEFINITION AB002449
ACCESSION AB002449.1 GI:2943812
VERSION
KEYWORDS advanced lung cancer.
SOURCE Homo sapiens tissue_l1b:fetal brain CDNA to mRNA, clone:843Ex.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE Ueno,K.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-1997) to the DDBJ/EMBL/GenBank databases.
Kiyonobu Ueno, Osaka University Medical School, Department of
Medicine III, 2-2, Yamada-oka, Suita, Osaka, Japan, Suita, Osaka
565, Japan (E-mail:ueno-kek@nchu.hosp.90.jp, Tel: 06-879-3835,
Fax: 06-879-3839)
2 (sites)
REFERENCE Ueno,K., Kumagai,T., Kijima,T., Kishimoto,T. and Hosoe,S.
TITLE Cloning and tissue expression of CDNA's from chromosome 5q21-22
AUTHORS which is frequently deleted in advanced lung cancer
JOURNAL Hum. Genet. 102 (1), 63-68 (1998)
MEDLINE 98141129
FEATURES
source
location/Qualifiers
1..1228
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="843Ex"
/map="5q21-22"
/tissue_l1b="fetal brain"
<1..>1228
/note="expressed ubiquitously"

BASE COUNT 415 a 225 c 205 g 377 t 6 others
ORIGIN
Query Match 5.4%; Score 56; DB 9; Length 1228;
Best Local Similarity 100.0%; Pred. No. 7.9e-21;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tggcgccgctctagaactagtgatcccccggcgctgcagaatcggcagcagct 56
|||||
DB 31 TGGCGCGCGCTCTAGAACTAGTGATCCCGCGGCTGCAGAAATTCGCGACGACT 86
RESULT 12
AX101122 1388 bp DNA PAT 10-APR-2001
LOCUS Sequence 10 from Patent WO0121650.
DEFINITION AX101122
ACCESSION AX101122
VERSION AX101122.1 GI:13619973
KEYWORDS
SOURCE
ORGANISM Dimorphotheca.
Dimorphotheca
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids II; Asterales; Asteraceae; Asteroideae;
Celastrales.
1 (bases 1 to 1388)
REFERENCE Coldren,C., Flint,D., Hallahan,D.L. and Wang,H.
AUTHORS Cis-prenyltransferases from plants
TITLE Patent: WO 0121650-A 1 29-MAR-2001;
JOURNAL E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
source
location/Qualifiers
1..1388
/organism="Dimorphotheca"
/db_xref="taxon:13490"
BASE COUNT 437 a 268 c 297 g 386 t
ORIGIN
Query Match 5.4%; Score 56; DB 6; Length 1388;
Best Local Similarity 100.0%; Pred. No. 7.8e-21;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tggcgccgctctagaactagtgatcccccggcgctgcagaatcggcagcagct 56
|||||
DB 221 TGGCGCGCGCTCTAGAACTAGTGATCCCGCGGCTGCAGAAATTCGCGACGACT 286
RESULT 13
114842 1601 bp DNA PAT 02-APR-1996
LOCUS Sequence 10 from patent US 5455167.
DEFINITION 114842
ACCESSION 114842
VERSION 114842.1 GI:1249751
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1601)
AUTHORS Voelker,T.A. and Davies,H.M.
TITLE Medium-chain thioesterases in plants
JOURNAL Patent: US 5455167-A 10 03-OCT-1995;
MEDLINE
FEATURES
source
location/Qualifiers
1..1601
/organism="unknown"
BASE COUNT 375 a 372 c 429 g 425 t
ORIGIN
Query Match 5.4%; Score 56; DB 6; Length 1601;
Best Local Similarity 100.0%; Pred. No. 7.7e-21;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

misc_feature      23455..23821
                  /note="Single clone region. Assembly confirmed by
                  restriction digest data."
misc_feature      65019..65283
                  /note="Single clone region. Sequence from reads from a
                  short insert library derived from a single pUC clone.
                  Restriction digest data confirm the assembly."
misc_feature      225123..225370
                  /note="Single clone region. Sequence from reads from a
                  short insert library derived from a single pUC clone.
                  Restriction digest data confirm the assembly."
BASE COUNT      56329 a 57092 c 57267 g 54682 t
ORIGIN

```

Query Match	25.5%	Score 463	DB 9	Length 225370
Best Local Similarity	100.0%	Pred. No. 8e-233		
Matches 463	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1234	ggaagctcctaagcccatatgagacgtctctgctctgctcgcgtcttctgcatctcagtc			
Db 62905	GGAGGCTCTCAGGCCCTCATGAGACCTTTCGTTGCTTCCGCTTTCGATCTTCAGTGT			
QY 1294	cccttccacttctgtctgtgtccctgaactaaaggaagaccttgaacaataatcaagc			
Db 62965	CCCTTTCACCTTTCCTGTGTGCTCCCTGAACCTAAAGAAAGACTCTGGAACAATTCACAC			
QY 1354	ccattttagggggcgatgacagccatcactgagggatggaagcaagcctgttactccaag			
Db 63025	CCATTTTGAGGGCGATGACAGCCACTACTAGGGATGAGACAAGCCTGTACTGCCAAG			
QY 1414	ctgggcccacagcccaagaccctctgctctgcccaggaggacccaatccagccctcgag			
Db 63085	CTGGGCCCAAGCCAGAGCCCTCTGCTGCCCCCAAGGGAGCCGAAATCCAGCCCTTGGAG			
QY 1474	ctttggtctcgaaaggtccctctctctctgtcaatgctccctccagcccatgaacccgggta			
Db 63145	CCTTGGTCTGCAGGGTCCCTCTCTCTGTCATGTCCTCCAGCCCATGACCCGGGGCTA			
QY 1534	ggaagctcactgcccctccctggttccaaactctctgctctgtctcttgaagactcaagaactt			
Db 63205	GGAGGCTACTCTCCCTCTTCTTCACCTCTCTGCTGCTTGAAGACTCAAGAACACTT			
QY 1594	cgagctcttcagaaactgtcgatcagccctccatctgccaagactaaagcagggaagagag			
Db 63265	CGAGCTTTCAGAACTGGCGGTACAGCCCTTCCATGCGCAAGACTAAAGCGCGAGAGAGAG			
QY 1654	gtggagccctcaggaactctgtcttctctgtgctgtagaggtgtcttgg 1696			
Db 63325	GTGGCCCTCTAGGATCTTGTCTCTCTGCTGGAGTGTCTTTTG 63367			

RESULT	8			
G20347/c				
LOCUS	G20347	120 bp	DNA	
DEFINITION	human STS A005N15,	sequence tagged site.	STS	
ACCESSION	G20347			24-JUL-1996
VERSION	G20347.1	GI:1340684		
KEYWORDS	STS; STS sequence;	primer; sequence tagged site.		
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
AUTHORS	Adams,M.D.			
JOURNALS				
COMMENT	Unpublished (1996)			

Contact: Mark Adams
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Email: mdadams@igf.org

Primer A: AAGCTTTTATTGGCAAC
Primer B: CCTTAGGATTTTGCTTC
SIS size: 120
PCR Profile:

Denaturation:	96C	5min
Anneal:	54C	30sec
Extend:	72C	30sec
Denature:	95C	30sec
FinalExtend:	72C	5min
Cycles:	30	

GenomicDNA:	25 ng
Primer:	0.43 μ m each
dNTPs:	230 μ m each
AmpliTaq:	0.5 units
TagStart Ab:	0.5 units
Total Volume:	10 μ l

Buffer:

Tris-HCl	pH 8.8:	100	mM
KCl:		500	mM
MgCl ₂ :		20	mM
Triton X-100:		1%	
Concentration:		10X	

Prepared with primer pairs derived from THC89870: GenBank Accession Numbers- R17700, T59518, T59564.

FEATURES
SOURCE

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
    sts
        primer_bind
            primer_bind
                complement(101..120)
                    BASE COUNT
                        39 a
                        36 c
                        26 g
                        19 t
                        ORIGIN

```

Query Match 4.4%; Score 79; DB 11; Length 120;
 Best Local Similarity 100.0%; Pred. NO. 6e-30;
 Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY	1698	AGGTGGTCTCGGcaatcgcgcgtcctcacaagcgcgcctcctalccggaaagaat	1755
Db	81	AGGTGGGCTCTGGCGCTTGGTCGCTCCTCTACGCGGCTCCTTATCGGGAAGAAT	22
QY	1758	ttgtttgccaataaagaac	1776
Db	21	TTGTTTCCCAATAAAGAC	3

PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
PT e.g. Ischemia and diabetes -
XX
XX
XX Claim 11: Page 71-72; 124pp; English.
CC The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and
CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischaemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is rat GLUTX1.
XX
XX
SQ Sequence 478 AA:
Query Match 70.5%; Score 627; DB 22; Length 478;
Best Local Similarity 100.0%; Pred. No. 3.1e-64;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 LVPISAPADVHLGLAMLVAGSMCLFIAGFAGWGPPIPLMSEIFPLHIKGYATGCVL 78
DB 357 lvpisapadvhlglaqlavagsmclfiagfagwgppllmseifplhikgyatgcvl 416
QY 79 TNNFMALVYTKEPNSIMEILRPYGAFWLTAAFCILSVLFTLPFVPTKGTLEQITAH 136
DB 417 tnnfmalfvltkfnsmellrpygafwlttafcilsvlftltpvptkgtlreqitah 474
RESULT 2
AAB66939
ID AAB66939 standard; Protein: 478 AA.
XX
AC AAB66939;
XX
DT 17-APR-2001 (first entry)
XX
DE GLUTX1 consensus sequence.
XX
KM GLUTX; gene therapy: vaccine; hexose transport modulator; human; rat;
KM hexose transport disorder; ischaemia; diabetes; hyperglycaemia; murine;
KM hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
OS Homo sapiens.
OS Mus sp.
OS Rattus sp.
XX
PN WO200104145-A2.
XX
PD 18-JAN-2001.
XX
PF 14-JUL-2000; 2000WO-1B01042.
XX
PR 14-JUL-1999; 99US-0143907.
PR 27-AUG-1999; 99US-0151140.
PR 23-FEB-2000; 2000US-0184285.
PR 13-JUL-2000; 2000US-0616132.
XX
PA (UYLA-) UNIV LAUSANNE.
XX
PI Thorens B, Idberson M, Uldry M;
XX
DR WPI; 2001-112615/12.
XX
PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
PT e.g. Ischemia and diabetes -
XX
XX Claim 11: Page 74-75; 124pp; English.
XX
CC The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and
CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative

CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischaemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is a consensus sequence
CC for GLUTX1.
XX
XX
SQ Sequence 478 AA:
Query Match 67.3%; Score 598; DB 22; Length 478;
Best Local Similarity 93.2%; Pred. No. 7.2e-61;
Matches 110; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 19 LVPISAPADVHLGLAMLVAGSMCLFIAGFAGWGPPIPLMSEIFPLHIKGYATGCVL 78
DB 357 lvpisapadvhlglaqlavagsmclfiagfagwgppllmseifplhikgyatgcvl 416
QY 79 TNNFMALVYTKEPNSIMEILRPYGAFWLTAAFCILSVLFTLPFVPTKGTLEQITAH 136
DB 417 tnnfmalfvltkfnsmellrpygafwlttafcilsvlftltpvptkgtlreqitah 474
RESULT 3
AAB66934
ID AAB66934 standard; Protein: 477 AA.
XX
AC AAB66934;
XX
DT 17-APR-2001 (first entry)
XX
DE Murine GLUTX1.
XX
KM Murine; GLUTX; gene therapy: vaccine; hexose transport modulator;
KM hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
KM hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
OS Mus sp.
XX
PN WO200104145-A2.
XX
PD 18-JAN-2001.
XX
PF 14-JUL-2000; 2000WO-1B01042.
XX
PR 14-JUL-1999; 99US-0143907.
PR 27-AUG-1999; 99US-0151140.
PR 23-FEB-2000; 2000US-0184285.
PR 13-JUL-2000; 2000US-0616132.
XX
PA (UYLA-) UNIV LAUSANNE.
XX
PI Thorens B, Idberson M, Uldry M;
XX
DR WPI; 2001-112615/12.
DR N-PSDB; AAF55867.
XX
PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
PT e.g. Ischemia and diabetes -
XX
XX Claim 11: Page 73-74; 124pp; English.
XX
CC The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and
CC AAB66932-AAB66941). The GLUTX proteins are related to the facilitative
CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischaemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is murine GLUTX1.
XX
XX
SQ Sequence 477 AA:

PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 the prevention, diagnosis and treatment of hexose transport disorders,
 e.g. ischemia and diabetes -
 XX
 PS Claim 11, Page 71-72; 124pp; English.
 XX
 CC The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and
 CC AAB66932-AAB6941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is rat GLUTX1.
 CC
 XX
 S0 Sequence 478 AA:

Query Match 100.0%; Score 508; DB 22; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1.3e-55;
 Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EPADVHGLAMLVAGSMCLFTAGFVAGVGPVPLMLSEIFPLHKGATGVCVLTNMPMA 60
 Db 363 epadvhglamlvagsmclftagfavgvpgipwlmseifplhkvatgvcvltnwma 422
 OY 61 FLVTKKEFNSIMEILRPYGAFWLTAAFCILSVLETF 94
 Db 423 flvckefnsimeilrpygafwltaafcilsvlft 456

RESULT 2

AAB66939 ID AAB66939 standard; Protein: 478 AA.

XX AAB66939;

DT 17-APR-2001 (first entry)

XX GLUTX1 consensus sequence.

XX GLUTX: gene therapy; vaccine; hexose transport modulator; human; rat;

KM hexose transport disorder; ischemia; diabetes; hyperglycaemia; murine;

KM hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.

XX Homo sapiens.

XX Mus sp.

XX Rattus sp.

XX WO200104145-A2.

XX 18-JAN-2001.

XX 14-JUL-2000; 2000WO-IB01042.

XX 14-JUL-1999; 9905-0143907.

XX 27-AUG-1999; 9905-0151140.

XX 23-FEB-2000; 2000US-0184285.

XX 13-JUL-2000; 2000US-0616132.

XX (UYLA-) UNIV LAUSANNE.

XX Thorens B, Ibberson M, Uldry M;

XX WPI: 2001-112615/12.

XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 the prevention, diagnosis and treatment of hexose transport disorders,
 e.g. ischemia and diabetes -
 XX
 PS Claim 11, Page 74-75; 124pp; English.

CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is a consensus sequence
 CC for GLUTX1.
 CC
 XX
 S0 Sequence 478 AA:

Query Match 95.9%; Score 487; DB 22; Length 478;
 Best Local Similarity 92.6%; Pred. No. 5.6e-53;
 Matches 87; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 EPADVHGLAMLVAGSMCLFTAGFVAGVGPVPLMLSEIFPLHKGATGVCVLTNMPMA 60
 Db 363 epadvhglamlvagsmclftagfavgvpgipwlmseifplhkvatgvcvltnwma 422
 OY 61 FLVTKKEFNSIMEILRPYGAFWLTAAFCILSVLETF 94
 Db 423 flvckefnsimeilrpygafwltaafcilsvlft 456

RESULT 3

AAB66934 ID AAB66934 standard; Protein: 477 AA.

XX AAB66934;

DT 17-APR-2001 (first entry)

XX Murine GLUTX1.

XX Murine: GLUTX: gene therapy; vaccine; hexose transport modulator;

KM hexose transport disorder; ischemia; diabetes; hyperglycaemia;

KM hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.

XX Mus sp.

XX WO200104145-A2.

XX 18-JAN-2001.

XX 14-JUL-2000; 2000WO-IB01042.

XX 14-JUL-1999; 9905-0143907.

XX 27-AUG-1999; 9905-0151140.

XX 23-FEB-2000; 2000US-0184285.

XX 13-JUL-2000; 2000US-0616132.

XX (UYLA-) UNIV LAUSANNE.

XX Thorens B, Ibberson M, Uldry M;

XX WPI: 2001-112615/12.

XX N-PsDB: AAF55867.

XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in
 the prevention, diagnosis and treatment of hexose transport disorders,
 e.g. ischemia and diabetes -
 XX
 PS Claim 11, Page 73-74; 124pp; English.

XX The present invention relates to GLUTX proteins (AAFS5865-AAFS5871 and
 CC AAB66932-AAB6941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is murine GLUTX1.
 CC
 XX
 S0 Sequence 477 AA:

